

A Measurement of the Top Quark Mass in 1.96 TeV Proton-Antiproton  
Collisions Using a Novel Matrix Element Method

by

John Freeman

B.A. (Columbia University) 2000

M.A. (University Of California, Berkeley) 2002

A thesis submitted in partial satisfaction of the  
requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Divison

of the

University of California, Berkeley

Committee in Charge:

Professor James Siegrist, Chair

Professor Robert Jacobsen

Professor Eugene Haller

Fall 2007



The dissertation of John Freeman is approved:

Chair	Date
	Date
	Date





## Abstract

### A Measurement of the Top Quark Mass in 1.96 TeV Proton-Antiproton Collisions Using a Novel Matrix Element Method

by

John Freeman

Doctor of Philosophy in Physics

University of California, Berkeley

Professor James Siegrist, Chair

A measurement of the top quark mass in  $t\bar{t} \rightarrow l + \text{jets}$  candidate events, obtained from  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV at the Fermilab Tevatron using the CDF II detector, is presented. The measurement approach is that of a matrix element method. For each candidate event, a two dimensional likelihood is calculated in the top pole mass and a constant scale factor, “JES”, where JES multiplies the input particle jet momenta and is designed to account for the systematic uncertainty of the jet momentum reconstruction. As with all matrix element techniques, the method involves an integration using the Standard Model matrix element for  $t\bar{t}$  production and decay. However, the technique presented is unique in that the matrix element is modified to compensate for kinematic assumptions which are made to reduce computation time. Background events are dealt with through use of an event observable which distinguishes signal from background, as well as through a cut on the value of an event’s maximum likelihood. Results are based on a  $955 \text{ pb}^{-1}$  data sample, using events with a high- $p_T$  lepton and exactly four high-energy jets, at least one of which is tagged as coming from a  $b$  quark; 149 events pass all the selection requirements. We find  $M_{meas} = 169.8 \pm 2.3 \text{ (stat.)} \pm 1.4 \text{ (syst.) GeV}/c^2$ .

---

Chair

Date





*To Tania, for all she gave.*

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>The Physics</b>	<b>7</b>
2.1	The Standard Model . . . . .	8
2.1.1	The Electroweak Force . . . . .	9
2.1.2	The Strong Force . . . . .	11
2.1.3	The Higgs Boson . . . . .	12
2.2	Beyond the Standard Model . . . . .	13
2.3	The Top Quark . . . . .	15
2.3.1	Production . . . . .	15
2.3.2	Decay . . . . .	16
2.3.3	Top Mass in Constrained Fits . . . . .	19
<b>3</b>	<b>Experimental Apparatus</b>	<b>25</b>
3.1	Overview . . . . .	26
3.2	The Accelerator Complex . . . . .	26
3.2.1	Overview of Luminosity . . . . .	26
3.2.2	Accelerating Protons before Tevatron Entry . . . . .	27
3.2.3	Creating and Accelerating Antiprotons . . . . .	28
3.2.4	The Main Injector and Recycler . . . . .	30
3.3	The CDF Detector . . . . .	30

3.3.1	Geometry . . . . .	32
3.3.2	Tracking . . . . .	33
3.3.3	Solenoid . . . . .	35
3.3.4	Calorimetry . . . . .	36
3.3.5	Muon System . . . . .	39
3.4	Trigger System . . . . .	41
3.4.1	Level 1 Trigger . . . . .	42
3.4.2	Level 2 Trigger . . . . .	43
3.4.3	Level 3 Trigger . . . . .	43
<b>4</b>	<b>Event Reconstruction and Selection</b>	<b>45</b>
4.1	Overview . . . . .	46
4.2	Electron Reconstruction . . . . .	47
4.3	Muon Reconstruction . . . . .	48
4.4	Jet Reconstruction . . . . .	48
4.4.1	Jet Clustering and the Raw Measurement . . . . .	50
4.4.2	The Relative Correction . . . . .	50
4.4.3	Multiple Interactions Correction . . . . .	51
4.4.4	Absolute Correction . . . . .	51
4.4.5	$E_T$ Selection Requirement . . . . .	51
4.4.6	Missing $E_T$ ( $\cancel{E}_T$ ) . . . . .	52
4.5	Secondary Vertex $b$ -tagging . . . . .	53
4.6	Data, Monte Carlo, and Background . . . . .	54
4.6.1	Data . . . . .	54
4.6.2	Monte Carlo . . . . .	54
4.6.3	Background . . . . .	56
<b>5</b>	<b>Signal Likelihood</b>	<b>61</b>
5.1	The Full Formula . . . . .	62

5.2	Comments on the Technique . . . . .	64
5.3	Phase Space Integration . . . . .	66
5.4	The Matrix Elements . . . . .	69
5.5	$M(M_t, \vec{x}) \rightarrow M_{eff}(M_t, \vec{x})$ : The Effective Propagators . . . . .	70
5.5.1	Construction of Effective Events . . . . .	71
5.5.2	Constructing the Effective Propagators . . . . .	74
5.6	Transfer Functions . . . . .	75
5.6.1	Overview . . . . .	75
5.6.2	Construction . . . . .	76
5.6.3	Transfer Function Study . . . . .	77
5.7	Normalization . . . . .	80
5.7.1	Normalization Without Effective Propagator Modification . . . . .	81
5.7.2	Addition of the Effective Propagator Normalization . . . . .	82
5.8	Acceptance . . . . .	85
5.9	Permutation weighting . . . . .	86
5.10	Integration details . . . . .	88
<b>6</b>	<b>Undesirable Events</b>	<b>92</b>
6.1	Introduction . . . . .	93
6.2	Signal Likelihood Cut . . . . .	94
6.3	Background Handling . . . . .	95
6.3.1	Modified Event Likelihood . . . . .	95
6.3.2	Calculation of a Background Probability, $f_{bg}$ . . . . .	99
<b>7</b>	<b>Tests of the Method</b>	<b>107</b>
7.1	Performing a Measurement . . . . .	108
7.2	Overview of Pseudoexperiments . . . . .	108
7.3	Signal-Only PEs . . . . .	110
7.4	Fully Realistic PEs . . . . .	111

7.5	Analysis Study . . . . .	115
<b>8</b>	<b>The Measurement</b>	<b>119</b>
8.1	The Measurement . . . . .	120
8.2	Measurement Crosschecks . . . . .	122
<b>9</b>	<b>Systematics</b>	<b>125</b>
9.1	Overview . . . . .	126
9.2	Individual Systematics . . . . .	127
9.2.1	Jet Energy Scale . . . . .	127
9.2.2	$b$ Jet Energy Scale . . . . .	128
9.2.3	MC Type . . . . .	129
9.2.4	ISR/FSR . . . . .	130
9.2.5	PDFs . . . . .	130
9.2.6	$gg$ Fraction . . . . .	132
9.2.7	Lepton $p_T$ . . . . .	133
9.2.8	Multiple Interactions . . . . .	133
9.2.9	$b$ -Tag $E_T$ Dependence . . . . .	134
9.2.10	Background Systematics . . . . .	134
9.2.11	Analysis-Specific Systematics . . . . .	136
9.3	A Comment on 2-d vs. 1-d Likelihood . . . . .	137
<b>10</b>	<b>Conclusions</b>	<b>140</b>
10.1	Conclusions . . . . .	141
<b>A</b>	<b>Data vs. MC Comparison</b>	<b>148</b>
<b>B</b>	<b>Details of the Kleiss-Stirling Matrix Elements</b>	<b>150</b>
B.0.1	Quark Annihilation Matrix Element . . . . .	151
B.0.2	Gluon-Gluon Fusion Matrix Element . . . . .	152

<b>C</b>	<b>Johnson’s Functions</b>	<b>156</b>
C.1	Johnson’s Functions . . . . .	156
C.2	Fitting . . . . .	157
<b>D</b>	<b>Transfer Function Crosscheck</b>	<b>159</b>
D.0.1	$W$ Mass Likelihood Derivation . . . . .	159
D.0.2	Top Mass Likelihood Derivation . . . . .	161
D.0.3	Priors . . . . .	162

# List of Figures

2.1	<i>Feynman diagrams for <math>t\bar{t}</math> production via <math>q\bar{q}</math> annihilation and gluon-gluon fusion. Note that in the rightmost diagram, the incoming gluons can switch vertices. . . . .</i>	16
2.2	<i>Feynman diagram of <math>t\bar{t} \rightarrow l + \text{jets}</math> . . . . .</i>	18
2.3	<i>Measurements of the top quark mass at CDF as of March 2007. The CDF and the world averages are shown at the bottom. . . .</i>	20
2.4	<i>The <math>W</math> propagator, and hence its coupling to the Higgs' vacuum expectation value (i.e., <math>M_W</math>) is affected through internal loops with the top quark and the Higgs. . . . .</i>	21
2.5	<i>Central 68% confidence region of the world average top and <math>W</math> mass measurements, lying outside the contour of <math>M_H = 114 \text{ GeV}/c^2</math></i>	22
2.6	<i><math>M_H</math> vs. <math>\chi^2</math> of a Standard Model global fit to high-<math>q^2</math> measurements; the blue region is LEP-excluded. . . . .</i>	23
3.1	<i>A sketch of the chain of accelerators used to produce <math>\sqrt{s} = 1.96 \text{ TeV}</math> <math>p\bar{p}</math> collisions. . . . .</i>	28
3.2	<i>Total luminosity delivered by the Tevatron, as function of store #. Over <math>3 \text{ fb}^{-1}</math> has been delivered, with a goal through FY09 of <math>8 \text{ fb}^{-1}</math> . . . . .</i>	29
3.3	<i>Initial instantaneous luminosities, by store, during Tevatron's Run II. . . . .</i>	31

3.4	<i>Cross-sectional sketch of the CDF detector, from the CDF Run II Technical Design Report . . . . .</i>	32
3.5	<i>Cross-sectional view of silicon tracking system (L00, SVXII and ISL) . . . . .</i>	34
3.6	<i>Partial cross-section of the COT. Every superlayer is 12 sense wires deep, and is segmented into chambers, the number of which at each superlayer is written on the picture. . . . .</i>	35
3.7	<i>Coverage, in <math>\eta</math>-<math>\phi</math> space, of the muon system. . . . .</i>	40
3.8	<i>Lower chambers for the Central Muon Extension and Scintillators, known as the “miniskirt” and possessing a planar fan geometry. . . . .</i>	41
3.9	<i>Dataflow of the CDF trigger system. The acceptance rate is drastically lower than the bunch collision rate. . . . .</i>	42
4.1	<i>Systematic errors on jet energy corrections as function of jet <math>p_T</math>. . . . .</i>	49
4.2	<i>A view of a <math>b</math> jet in the x-y plane. The primary means of determining whether a jet originates from a heavy flavor quark is through the value of <math>L_{2D}</math>. . . . .</i>	54
5.1	<i>Profile plots of differences between measured jet and true parton angles for <math>b</math> jets (left) and light jets (right), <math>\phi</math> (top) and <math>\eta</math> (bottom), as a function of jet <math>p_T</math>. . . . .</i>	73
5.2	<i>The hadronic-side and leptonic-side propagators in <math>m_W^2</math> and <math>m_t - M_t</math>. In the integration, <math>m_t - M_t</math> is converted to <math>m_t^2</math> for an assumed <math>M_t</math>. . . . .</i>	75
5.3	<i>Fitted transfer functions for light quarks in the fiducial region of the central calorimeters (<math>0.15 &lt;  \eta  &lt; 0.85</math>). The solid line shows the region of the distribution to which the function was fitted, while the dotted line represents the region of extrapolation. . . . .</i>	78



5.4	<i>Transfer function cross check results for hadronic <math>W</math> and top mass using fully simulated Herwig MC.</i>	81
5.5	<i>Our normalization, before the modification for the effective propagators, compared to the Herwig cross section after an appropriate scaling.</i>	82
5.6	<i>The normalization correction applied to account for our use of effective propagators.</i>	84
5.7	<i>The acceptance applied to the denominator in the likelihood function.</i>	87
5.8	<i>5<sup>th</sup> degree polynomial fits to <math>b</math> tag efficiencies as a function of <math>\eta</math> and <math>p_T</math> for both <math>b</math> quarks and light quarks.</i>	88
5.9	<i>The points of the <math>t\bar{t}</math> transverse momentum integration grid, chosen to reflect the values found in MC.</i>	90
6.1	<i>Distribution of maximum log likelihood values for good signal (red), bad signal (green) and background (blue) in the 1-tag (top) and &gt;1-tag (bottom) cases.</i>	96
6.2	<i>Distributions of the discrimination variable, <math>q</math>; the histograms for signal and backgrounds are scaled to reflect the expected fractions of signal/backgrounds in the sample.</i>	100
6.3	<i>Distributions of aplanarity, <math>H_{TZ}</math> and <math>D_R</math>. Signal is at <math>M_t = 180</math> GeV/<math>c^2</math>.</i>	103
6.4	<i>Bias of mean (black markers) and RMS (white markers) of distributions of <math>A</math>, <math>H_{TZ}</math> and <math>D_R</math> for signal MC at different masses with respect to the mean and RMS of the distributions at <math>M_t = 180</math> GeV/<math>c^2</math>. The solid lines are fits to the biases on the mean, the dotted lines represent the bias on the mean of the two types of background under study.</i>	104

6.5	With the same format of the plots shown in Fig. 6.4, the plots here show the effect of shifting the jet energy corrections by + and - $1\sigma$ . . . . .	105
7.1	Results of PEs run with 179 good signal events each. Clockwise, from upper left: $\langle M_{meas} \rangle$ vs. $M_t$ , $\langle M_{meas} \rangle - M_t$ vs. $M_t$ , pull width vs. $M_t$ , and RMS of $M_{meas}$ vs. $M_t$ . . . . .	112
7.2	Results of PEs run with 179 signal events, good and bad. Plot format same as in Fig. 7.1 . . . . .	113
7.3	The results of fully realistic PEs. Plot format same as in Fig. 7.1.	114
7.4	Results of the JES study. Top left, $\langle M_{meas} \rangle$ vs. $M_t$ for different JES values. Top right, $\langle JES_{meas} \rangle$ vs. JES for different $M_t$ values. Bottom, $\langle M_{meas} \rangle$ vs. JES for different $M_t$ values. . .	116
8.1	Total 2-d log likelihood of the 149 events in the data sample. . .	120
8.2	Positive and negative errors of PEs run with signal MC at $M_t = 170.0 \text{ GeV}/c^2$ . The red arrows indicate the positive and negative errors of the data measurement. . . . .	121
8.3	Mass and log likelihood values at individual event likelihood maxima. Black markers are data, blue histograms are signal MC at $M_t = 170.0 \text{ GeV}/c^2$ and red histograms are background; signal and background are combined in their expected proportions of the data sample. Plot of masses includes likelihood cut, plot of maximum log likelihoods excludes the cut. . . . .	123
9.1	Effect of using different PDFs on the reconstructed top mass. .	132
A.1	Comparison of data vs. MC, jet $E_T$ 's. . . . .	149
A.2	Comparison of data vs. MC, various event observables. . . . .	149

D.1	<i>2-d distributions of <math>\cos \theta</math> and <math>p</math> taken from MC at the generator level for both the hadronic-side <math>W</math> and top.</i>	164
D.2	<i>1-d functions used for the particle priors, fit to generator-level quantities.</i>	164



# List of Tables

2.1	<i>The fermions of the Standard Model. There is an immense spread of masses among particles across the three generations [12]. The masses are given here in units of <math>\text{GeV}/c^2</math>, to three significant digits, where possible, and without quoted errors. <math>\nu</math> masses are not measured directly, though known to be much smaller than other particle masses but nonzero.</i>	9
3.1	<i>Statistics for the calorimeters.</i>	37
4.1	<i>The three datasets, corresponding to different periods of Tevatron running, used in our measurement.</i>	55
4.2	<i>Expected number of background events given the selection cuts described in this chapter. Events with more than 1 <math>b</math> tag have a much lower proportion of background than events with only 1 <math>b</math> tag.</i>	57
5.1	<i>Here we have the # of points which sample the given variable's dimension, and info on whether the spacing of these points is equidistant in probability or constant in the variable</i>	90
6.1	<i>Efficiency of the selection cut <math>\log L_{max} &gt; 6</math> at <math>M_t = 175 \text{ GeV}/c^2</math></i>	95
6.2	<i>Expected numbers of background and signal in data after likelihood cut.</i>	95

7.1	<i>PE results with the log likelihood cut and/or the background handling unused, as well as the fully realistic case in which both are used.</i>	117
8.1	<i>Top mass measurement summary.</i>	123
9.1	<i>Total list of systematics.</i>	128
9.2	<i>Results for jet energy uncertainty.</i>	129
9.3	<i>Systematics from ISR/FSR. All samples have a nominal <math>M_t</math> of 178 GeV.</i>	130
9.4	<i>PDF crosscheck results</i>	132
9.5	<i>Study of reconstructed mass as function of number of vertices in Herwig sample at <math>M_t = 175 \text{ GeV}/c^2</math>.</i>	134
9.6	<i>Background systematics.</i>	136
9.7	<i><math>\overline{\log L_{bg}(M_t, \text{JES})}</math> systematic</i>	137



## Acknowledgements

The number of people who have positively affected my life during my time in graduate school are too numerous to mention in just a page or two, and some omissions must inevitably exist in these acknowledgements. However, many individuals stand out in my mind, and will be mentioned here.

At the various stages of my graduate career, I have had the pleasure of working with many excellent physicists. Very special praise goes to my advisor, Jim Siegrist, who, despite having a schedule similar to that of the Secretary General of the UN, nonetheless always had time for me and my "updates".

Early in graduate school, during my work on the ATLAS pixel detector, Kevin Einsweiler, John Richardson and Maurice Garcia-Sciveres all provided helpful guidance. A special thanks goes to the then newly-minted postdoc, Aldo Saavedra, who was not only a great supervisor of mine but also put up with my various shenanigans with good humor and forbearance. Another great supervisor was Marc Weber, with whom I worked on the CDF Run IIb silicon detector.

Thanks also goes out to all members of the LBNL group who worked on the method described in this thesis: Jason Nielsen, Jeremy Lys, Lina Galtieri, Igor Volobouev, Pedro Movilla Fernandez and Paul Lujan. Best of luck to all of them as this analysis proceeds into the future!

Additional thanks go to those at CDF who guided this analysis through its blessing, including the top mass conveners, Daniel Whiteson and Florencia Canelli, the top group conveners, Robin Erbacher and Kirsten Tollefson, and the internal reviewers, George Velez, Jahred Adelman, and Ricardo Eusebi. They all gave a great deal of their time and attention to this analysis, and I appreciate their efforts.

My thesis committee deserves great credit for their efforts during my PhD approval process, especially due to their reading of this document despite very



busy schedules. This committee is composed of Jim, Bob Jacobsen, and Eugene Haller.

Additional acknowledgement goes to Weiming Yao and Doug Glenzinski. Doug Glenzinski graciously wrote me a recommendation, and Weiming was a fountain of good advice during the last few years of my graduate school experience.

Finally, I give thanks to those outside my professional life. Thanks to Rob Lillis, my liason to every social event in existence at Berkeley. Thanks to Patrick Dunn, a good friend. Thanks to Henri Bachacou, who laughed at my jokes and brightened my days at the lab. And thanks to Mike Grobis, a fellow ex-wrestler who accompanied me on my creatine-fueled adventures in the Berkeley gym. My acknowledgements also go to the denizens of the CDF students office: Hung-Chung Fang, Johannes Muelmenstaedt, and Amanda Deisher. Though Paul has already been mentioned, it is more than appropriate to register my appreciation for him here as well. And a very special thanks goes to Tania D'Alberti, a pillar of support for me during hard times and a ray of goodness in this world.

Most of all, I give thanks to Mom and Dad, and their limitless contribution to my life.



# Chapter 1

## Introduction

If physics can be considered the most fundamental of all the sciences, one can add that particle physics is the most fundamental branch within physics. A reductionist view of the universe maintains that all phenomena are the result of the interactions of fundamental particles through the four forces of gravity, electromagnetism, the weak nuclear and the strong nuclear force. With this in mind, the holy grail of particle physics is to find a so-called “Theory of Everything”, predicting both the spectrum of the fundamental particles and unifying the four forces as different manifestations of a single more fundamental interaction. However, as wonderful a theory as this would be, it should be noted that the phrase “Theory of Everything” is, of course, an exaggeration. While a complete knowledge of particles and their interactions provided by a Theory of Everything could explain all phenomena *in principle*, it is clear that the sheer number and complexity of these interactions would render it useless as a tool to, say, better predict the turbulent behavior of fluids, or indeed any complex macroscopic phenomenon involving many, many fundamental particle interactions.

Nevertheless, the goal of looking for the underlying unity in nature is a worthy one, and while a Theory of Everything may never be found, a great deal of knowledge can be developed in its pursuit. Indeed there *is* a great deal of fascinating knowledge contained within the best current theory of fundamental particles and their interactions, the Standard Model. Since the early 1970’s, the Standard Model has proved to be an astonishingly successful theory of fundamental particle interactions at the energy scales attainable by accelerators all the way up to the present-day Tevatron synchrotron at the Fermi National Accelerator Laboratory. Contained in the Standard Model is information on why the sun shines, why an atomic nucleus doesn’t fly apart due to the electromagnetic repulsion of protons, and why the universe would appear to obey different physical laws if viewed through a mirror. To date, all particles predicted by the

Standard Model have been found; notable examples are the massive  $W^\pm$  and  $Z^0$  gauge bosons posited in the Standard Model as carriers of the weak force, seen at CERN's Super Proton Synchrotron in 1983 [1, 2, 3, 4], and the top quark as the partner of the bottom quark in an  $SU(2)_L$  electroweak isospin doublet just over a decade ago at the Tevatron [5]. Indeed, the only remaining particle predicted by the Standard Model but as-yet undiscovered is the Higgs boson, a spin-0 particle required in the model to interact with particles in order to give them mass. As more data is collected during Run II at the Tevatron, finding evidence of a Standard Model-like Higgs boson at  $3\sigma$  (99.7%) significance is considered to be a possibility; its discovery (requiring a  $5\sigma$ , or 99.99994% significance) would have to wait for the CERN's Large Hadron Collider (LHC), scheduled to turn on next year (2008).

Despite the ability of the Standard Model to use the Higgs to describe how particles possess mass, however, it is unable to explain why the particles have such different masses. In the late 1970's, when the bottom quark was discovered and found to be the most massive quark to date [6], most physicists would not have guessed that the top quark consequently predicted by the Standard Model would turn out to be 35-40 times as massive as the bottom quark and would consequently take 18 years to discover! With the approximate mass of a gold nucleus, the top quark is the heaviest known fundamental particle.

The large mass of the top quark is of interest for a couple of reasons. One is that it has a large effect on radiative corrections to electroweak parameters within the Standard Model; as a result it can be used in conjunction with the measured mass of the  $W$  boson to provide a constraint on the mass of the elusive Higgs. Another is that it can provide constraints beyond the Standard Model; e.g. in the Minimal Supersymmetric Standard Model where it can provide constraints on the lightest neutral Higgs. Also the large mass of the top gives it a special role in extensions to the Standard Model, as it is kinematically

permitted to decay to non-Standard Model particles such as charged Higgs or stop quarks in supersymmetric theories. Finally, the top’s high mass makes it unique among the quarks in that it decays in a time an order of magnitude shorter than the characteristic timescale for hadronization through the strong force,  $\hbar/\Lambda_{QCD} \sim \mathcal{O}(10^{-24} - 10^{-23}s)$ . As a result of this, the top is essentially a “bare” quark, whose properties such as its mass and spin can be measured directly through its decay products.

Due to the features of the top described above, there is a very rich program at the Tevatron to measure various properties of the top quark. The measurement which has drawn the most attention since the discovery of the top has been that of its mass. This thesis concerns a measurement of the mass of the top quark developed at Lawrence Berkeley National Laboratory as part of the broader effort of CDF’s top mass measurement program, using 149 candidate events from  $955 \text{ pb}^{-1}$  of data collected over a period from February 2002 to February 2006. The basic technique employed is a “matrix element method”, conceived of in the late 1980’s [7] and first used to measure the top quark mass by D0 in their Run I dataset [8]. Matrix element methods involve an event-by-event calculation of a likelihood as a function of top quark pole mass,  $M_t$ ; their attractive features include a natural event-by-event weighting due to the width of the likelihood, the ability to model non-gaussian calorimeter energy resolution, and the use of Standard Model information in the form of an integration over the top-antitop ( $t\bar{t}$ ) production and decay matrix element. In recent years, as computational power has grown to match the heavy calculational demands of the matrix element approach, this method has become an extremely popular type of measurement, with numerous top mass measurements at CDF and D0 in different  $t\bar{t}$  decay channels [9, 10, 11].

The matrix element technique, as well as our unique implementation of it, will be comprehensively described in this thesis. However, the thesis will

begin by covering the relevant physics of interest underlying the top quark mass, touched on in this introduction. I will proceed with a description of the experimental apparatus, both the CDF detector itself and the series of Fermilab accelerators which culminate in the Tevatron to accelerate protons and antiprotons to a still-record high center-of-mass energy of 1.96 TeV. A discussion of how the various decay objects from the top are reconstructed in the detector will follow, along with a discussion of the selection cuts used to obtain the data sample for the measurement and the Monte Carlo events used in construction and evaluation of the analysis. The measurement technique itself will be described at length, followed by the measurement, a discussion of systematics, and a brief conclusion.

## Chapter 2

## The Physics



## 2.1 The Standard Model

To understand the physics of the top quark, it is helpful to first understand the top within the context of the Standard Model. The Standard Model is currently the most successful theory of fundamental particle interactions, describing spin-1/2 fermions interacting through spin-1 gauge bosons, all of whose masses come about due to interactions with a spin-0 boson known as the Higgs boson. Broadly speaking, the fermions can be considered “matter” particles which obey the Pauli exclusion principle, and the spin-1 bosons can be considered “force-carrying” particles, which do not. The model describes three of the four known forces, the strong, electromagnetic and weak; within it exists a (partial) unification of the electromagnetic and weak forces. Gravity has proved difficult to incorporate into quantum field theories due to problems with avoiding infinities in interaction probability calculations given the presence of non-flat spacetime; a further discussion is beyond the scope of this thesis. Fortunately for the experimentalist, the strength of gravity is many, many orders of magnitude weaker than even the weak force and is not a concern in experimental particle physics!

The Standard Model groups the fermions into three generations; differences between the generations concern only the masses of the particles rather than their allowed interactions. There are two types of fermions, leptons and quarks. The leptons consist of three negatively charged particles, the electron ( $e$ ), the muon ( $\mu$ ) and the tau ( $\tau$ ) as well as three neutral partners, the neutrinos,  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ . The quarks are either up-type, with a charge of  $+2/3$  e, or down-type, with a charge of  $-1/3$  e. The up-type quarks consist of the up ( $u$ ), the charm ( $c$ ) and the top ( $t$ ) while the down-type quarks consist of the down ( $d$ ), the strange ( $s$ ) and the bottom ( $b$ ).

In the Standard Model, every particle has a corresponding antiparticle of the same mass, but with opposite quantum numbers. It is interesting to note that

	1 <sup>st</sup> generation	2 <sup>nd</sup> generation	3 <sup>rd</sup> generation
charged leptons, $l$	$M_e = 0.511 \times 10^{-3}$	$M_\mu = 0.106$	$M_\tau = 1.78$
neutrinos, $\nu$	$M_{\nu_e} > 0$	$M_{\nu_\mu} > 0$	$M_{\nu_\tau} > 0$
up-type quarks	$M_u = 1.5 - 3.0 \times 10^{-3}$	$M_c = 1.25$	$M_t = 171$
down-type quarks	$M_d = 3 - 7 \times 10^{-3}$	$M_s = 0.095$	$M_b = 4.2 - 4.7$

Table 2.1: *The fermions of the Standard Model. There is an immense spread of masses among particles across the three generations [12]. The masses are given here in units of  $\text{GeV}/c^2$ , to three significant digits, where possible, and without quoted errors.  $\nu$  masses are not measured directly, though known to be much smaller than other particle masses but nonzero.*

despite the ostensible parallels between particles and antiparticles as well as across the three generations within the Standard Model, there are overwhelmingly more fermions (matter) than antifermions (antimatter) in the universe; furthermore, the vast majority of these fermions are electrons, up quarks and down quarks bound into protons and neutrons, and neutrinos. The fermions are distinguished from one another by their interactions: while all fermions experience the weak force, only quarks interact through the strong force, and only quarks and charged leptons experience the electromagnetic force. The Standard Model classifies the fermions in three generations, as shown in Table 2.1.

### 2.1.1 The Electroweak Force

Forces, transmitted by the spin-1 gauge bosons, are generated within the Standard Model Lagrangian by group symmetries of particle multiplets. The electromagnetic and weak forces become manifest from the breaking of the  $\text{SU}(2)_L \otimes \text{U}(1)_Y$  electroweak symmetry due to the Higgs boson (more on this below). The gauge boson of the electromagnetic field is the photon,  $\gamma$ , while the gauge bosons of the weak field are the  $W^\pm$  and  $Z^0$  bosons. While the  $\gamma$  is massless, the  $W$ 's and  $Z$ 's are very massive, 80.4 and 91.2  $\text{GeV}/c^2$ , respectively.

It is for this reason that the weak force seems “weak”. The amplitude at tree level for an interaction of four momentum  $q$  via a gauge boson of mass  $M_x$  with coupling  $x$  (and zero decay width) goes as  $(x^2)/(M_x^2 - q^2)$ . In the case of the photon, where  $M_\gamma^2 = 0$ , the amplitude for a low- $q^2$  interaction will be much larger (i.e., the force will be much stronger) than in the case of the  $W$  boson, as its mass squared will appear in the denominator of the amplitude, and the electromagnetic coupling,  $e$ , is the same order of magnitude as the weak coupling constant,  $g$ . The characteristic strength of a low energy weak interaction is described by Fermi’s constant,  $G_F = \frac{\sqrt{2}}{8}g^2/M_W^2$ ; its current measured value is  $1.16637(1) \times 10^{-5} \text{ GeV}^{-2}$ . On the other hand, for the electromagnetic field we have the fine structure constant,  $\alpha = e^2/(4\pi)$ , whose current measured value is  $1/137.03599911(46)$ .

As a result of having a massless boson, the electromagnetic field has an infinite range, while on the other hand, weak interactions occur over very short range,  $\sim \hbar c/M_W$ , or  $\sim 10^{-18} \text{ m}$ . Due to the infinite range of the electromagnetic field, we directly observe it in our everyday lives; on the other hand, the weak force is of great interest in that it is the only force which violates parity - i.e., one could tell whether or not one were observing weak interactions in a mirror. Additionally, it is capable of producing intergenerational quark flavor changes, as the physical mass eigenstates of quarks are superpositions of their flavor eigenstates, and hence, for example, a physical charm quark of the  $2^{nd}$  generation could emit a  $W$  and convert to a down quark of the 1st generation. The amplitudes involved in calculating the probability that a given physical quark will decay to a quark in a given generation are represented in the  $3 \times 3$  CKM (Cabbibo-Kobayashi-Maskawa) matrix:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.1)$$

where  $V_{ij}$  is a term in the amplitude for flavor changing weak interactions between up-type quark  $i$  and down-type quark  $j$ . Under the Standard Model's assumption of the unitarity of the CKM matrix, the current measured magnitudes of these complex terms are as follows [12]:

$$\begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272^{+0.0010}_{-0.0010} & (3.96^{+0.09}_{-0.09}) \times 10^{-3} \\ 0.2271^{+0.0010}_{-0.0010} & 0.97296^{+0.00024}_{-0.00024} & (42.21^{+0.10}_{-0.80}) \times 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \times 10^{-3} & (41.61^{+0.12}_{-0.78}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix} \quad (2.2)$$

As the values in the CKM matrix are complex, the unitarity constraint imposed by the Standard Model on the CKM matrix means it can be parameterized with four real numbers, and necessarily contains a CP violating phase. Searches for CP violation in weak interactions constitute an extremely rich vein of research in particle physics.

### 2.1.2 The Strong Force

The quanta of the strong field are massless gluons, which behave very differently than the other gauge bosons; the study of this field is known as “Quantum Chromodynamics”, or “QCD”. In the Standard Model Lagrangian the gluons are associated with an  $SU(3)_c$  group of rotations applied to color triplets of quarks, where “color” is the name of the strong charge, and a quark can possess a color charge of red, green or blue. The generators of the  $SU(3)_c$  group, however, are non-Abelian (i.e., do not commute); as a result the Standard Model correctly predicts that gluons may interact with one another, and hence themselves possess color charge. The consequence of this is that the coupling for the strong force,  $\alpha_s$ , increases with decreasing energy-momentum ( $q^2$ ) exchange between the gluons. This is the opposite behavior of the coupling in electromagnetism,  $\alpha$ , and in fact,  $\alpha_s$  runs as [13]

$$\alpha_s(q^2) = \frac{12\pi}{(33 - 2n_f(q^2)) \ln(\frac{q^2}{\Lambda_{QCD}^2})} \quad (2.3)$$

where  $\Lambda_{QCD}$  is an energy on the order of 0.2 GeV and  $n_f$  is the number of quark flavors which can be pair produced given  $q^2$ . Here it can be seen that for very high energy interactions ( $q^2 \gg \Lambda_{QCD}^2$ ),  $\alpha_s$  assumes a low value (i.e., the strong force isn't "strong" at all), but as the interaction energy decreases and approaches  $\Lambda_{QCD}$ ,  $\alpha_s$  grows very large. Therefore, for quarks interacting at distance scales  $\sim \hbar c / \Lambda_{QCD}$ , or  $\mathcal{O}(10^{-15}\text{m})$ , the interaction between them becomes so strong that quarks are compelled to be bound into color neutral configurations known as "hadrons". These are either the bosonic bound states of quark-antiquarks, known as "mesons", or the fermionic bound states of three quarks, known as "baryons". The most commonly known hadrons are the protons and neutrons of an atom's nucleus, whose diameters, as expected, are roughly about  $10^{-15}\text{m}$ . This property of the strong force is known as "asymptotic freedom." It is impossible to understate the impact this property has for the particle experimentalist, in that the kinematic properties of quarks in a decay can only be ascertained through the jet of particles into which they hadronize; this issue will be further discussed at various points in this thesis.

### 2.1.3 The Higgs Boson

All the particles in the Standard Model described so far have been discovered; indeed the only remaining particle yet to be discovered is the Higgs boson. The Higgs boson is needed as adding explicit mass terms to the Standard Model Lagrangian renders the theory non-renormalizable. Rather, it enters the theory as an  $SU(2)_L$  doublet of complex scalar fields:

$$\begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix} \quad (2.4)$$

In the Standard Model Lagrangian, the potential for these fields is expressed as

$$V(x) = \mu^2 |\Phi(x)|^2 + \lambda |\Phi(x)|^4 \quad (2.5)$$

where

$$|\Phi(x)|^2 = |\phi_1(x)|^2 + |\phi_2(x)|^2 + |\phi_3(x)|^2 + |\phi_4(x)|^2 \quad (2.6)$$

and if we have  $\mu^2 < 0$  and  $\lambda > 0$ , the potential will have a minimum at a nonzero field value. Three of the four fields in the doublet can be gauged away due to invariance of the Lagrangian under  $SU(2)$  rotations of the doublet; this leaves a single, final field whose particle is the Higgs, and whose value at the minimum potential is  $v \simeq 246$  GeV. It is this nonzero vacuum expectation value which gives particles with which the Higgs interacts nonzero mass and which breaks the electroweak  $SU(2)_L \otimes U(1)_Y$  symmetry; unfortunately, while the Higgs field provides a mechanism for how particles acquire mass, the Standard Model is incapable of predicting the particle masses. It is even incapable of explaining why particle masses are as varied as they are, despite the ostensible symmetries of the three generations shown in Table 2.1. This indicates that the Standard Model, for all its successes, is not a complete theory of fundamental particles and their interactions. Nevertheless, the search for the Higgs boson has been very enthusiastic within experimental particle physics, and is in fact perhaps the main justification for the construction of the LHC. As will be shown shortly, the measurement of the top quark mass plays an integral role in our current knowledge of the Higgs boson.

## 2.2 Beyond the Standard Model

Despite the success of the Standard Model, in recent years experimental results have been found which it had failed to predict. The discovery that

neutrinos have mass [14] must be incorporated into the Standard Model by the addition of a mixing matrix describing the difference between mass and flavor eigenstates; this matrix, the “PMNS” matrix, is the neutrino analog of the CKM matrix. Potentially less straightforward to model are cosmological observations of dark matter, composed of particles which are indirectly observable through, e.g., their gravitational effects on galaxy rotation rates and gravitational lensing effects. Dark matter’s calculated contribution to the total energy density in the universe of  $\sim 10\%-30\%$  is far more than that of ordinary baryonic matter which comprises only a few percent of the total. This dark matter may very likely be composed of particles not predicted by the Standard Model. Finally, dark energy, composing the remainder of the universe’s energy density, exerts a negative pressure which increases the rate of the expansion of the universe [15] and is also unexplained by the Standard Model.

Aside from the recent observations of physics that exists beyond the Standard Model, there have been longstanding concerns about its theoretical structure. One of the most serious is the so-called “hierarchy problem”. In the Standard Model, internal particle loops in the Higgs’ propagator radiatively correct the squared mass of the Higgs in a manner which diverges quadratically with increasing energies, all the way up to gravity’s energy scale of the Planck mass, which is about 16 orders of magnitude larger than the predicted electroweak scale of the Higgs mass,  $\mathcal{O}(100\text{GeV})$ . The Standard Model requires an *extremely* fine tuned balance of the radiative corrections provided by various particles, then, to allow a Standard Model Higgs mass. Other models of particle interaction, fortunately, eliminate the concern over the hierarchy problem. Supersymmetric theories accomplish this through postulation of bosonic “superpartner” particles to the fermions (and vice-versa) whose existence provide radiative corrections to the Higgs mass which balance out those of the “regular” particles and naturally allow for an electroweak-scale Higgs mass. Extra dimen-

sions, on the other hand, explain an electroweak-scale Higgs mass by removing concerns over the Planck scale through allowing gravity to weaken by its propagation into one or more extra dimensions inaccessible to other forces. While top mass measurement can constrain the mass of the Standard Model Higgs as will be described in Sec. 2.3.3, it is also capable of constraining the masses of neutral Higgs particles in supersymmetric extensions to the Standard Model.

## 2.3 The Top Quark

### 2.3.1 Production

The top quark is the heaviest known fundamental particle, with a mass equal to that of a gold nucleus; as of March 2007, the world average for its mass is  $170.9 \pm 1.8 \text{ GeV}/c^2$  [16]. As a consequence of this high mass, it was not until the 1990's, using the Tevatron at Fermilab during CDF's Run I (1992-1996) that sufficient collision energies were available to produce the top quark. This production was accomplished with  $\sqrt{s} = 1.8 \text{ TeV}$  center-of-mass collisions of protons and antiprotons; today it continues during CDF's Run II at  $\sqrt{s} = 1.96 \text{ TeV}$ . The Tevatron remains the only collider in the world capable of producing top quarks for study.

Even at the upgraded energies of Run II, the cross section for top quark production at the Tevatron is quite low. The top quark is produced in one of two ways: through the production of a single top quark, or, more commonly, through the production of a  $t\bar{t}$  pair. In practice, it is only events in which a  $t\bar{t}$  pair is created which are used to study properties of the top quark, with an exception noted below. There are a couple of reasons for this. One is that the theoretical Standard Model cross section for the production of a  $t\bar{t}$  pair ( $6.7 \pm 0.9 \text{ pb}^{-1}$  at  $M_t = 175 \text{ GeV}/c^2$  [17]) is more than twice as large as that of the production of a single top quark. More importantly, however, the backgrounds



for single top events are very high, such that while there is a rich program of top study at CDF using  $t\bar{t}$  pairs which will be discussed, evidence of single top events has only recently been found [18]. For the dataset of  $\mathcal{L}_{int} = 955 \text{ pb}^{-1}$  used in this thesis, then, it can be expected that several thousands of  $t\bar{t}$  pairs were produced. Nevertheless, the measurement uses a mere 149 events due to the stringent selection requirements placed on  $t\bar{t}$  candidate events, discussed in Chap. 4 and in Sec. 6.2.

The method by which  $t\bar{t}$  pairs are produced in a  $p\bar{p}$  collision is through  $q\bar{q}$  annihilation or gluon-gluon fusion, the Feynman diagrams for which are shown in Fig. 2.1. At next-to-leading order,  $q\bar{q}$  annihilation is calculated to account for 85% of the production cross section,  $\sigma_{t\bar{t}}$ , while gluon-gluon fusion only accounts for 15% of  $\sigma_{t\bar{t}}$  given the Tevatron collision energy of  $\sqrt{s} = 1.96 \text{ TeV}$ . There is some uncertainty on this, however; for gluon-gluon annihilation, calculations can provide a figure as low as 10%, or as as high as 20% [19]. It is interesting to note that at the LHC, gluon-gluon fusion is expected to produce a majority of  $t\bar{t}$  pairs, as the 7 TeV/proton energies make it far more probable that gluons carry enough energy for  $t\bar{t}$  production than at the Tevatron.

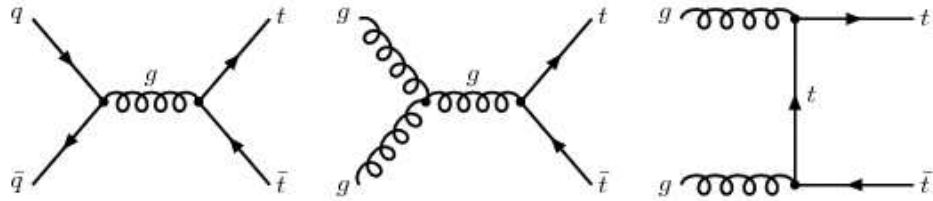


Figure 2.1: *Feynman diagrams for  $t\bar{t}$  production via  $q\bar{q}$  annihilation and gluon-gluon fusion. Note that in the rightmost diagram, the incoming gluons can switch vertices.*

### 2.3.2 Decay

In the Standard Model, the top quark almost always decays via the weak force to a  $W$  boson and a  $b$  quark. As can be seen in Eq. 2.2, assuming unitarity

of the CKM matrix, the amplitude for the  $t \rightarrow Wb$  decay is  $\propto V_{tb} = 0.9991$ . It is not known, however, if the 3-generation CKM matrix is in fact unitary, and measurements of  $V_{tb}$  are performed in analyses searching for single top events. These measurements are still very much statistics limited and errors on the  $V_{tb}$  measurement are currently on the order of 0.1-0.2 [18].

The timescale of the top decay is very brief due to the high mass of the quark,  $\tau = \hbar/\Gamma \simeq 0.5 \cdot 10^{-24}$ , where  $\Gamma$  is the calculated width of the top predicted in the Standard Model [20]. The top is unique among the quarks in that its decay occurs before it has a chance to hadronize, as the characteristic timescale of hadronization,  $\hbar/\Lambda_{QCD}$  is approximately an order of magnitude larger than that of the decay. The consequence of this is that the top can be studied as a “bare” quark; fundamental properties such as the top’s mass and spin properties can be measured directly through the top’s decay products before color confinement (i.e., hadronization) occurs.

After a top decays to a  $W$  and  $b$ , the  $b$  will convert into a  $B$  hadron, which will in turn decay into a jet of particles; the  $W$  will decay as well. The  $W$  can decay leptonically to an  $e$ ,  $\mu$  or  $\tau$  and their partner  $\nu$ ’s, or it can decay to a  $u\bar{d}$  or a  $c\bar{s}$  pair; the  $t\bar{b}$  pair, of course, is kinematically disfavored. Given that quarks carry three possible colors, the  $W$  decays to quarks twice as often as to leptons; this excludes the mass effects of the decay products on the branching ratios, which are minor as the  $W$ , at a mass of  $80.4 \text{ GeV}/c^2$ , is much more massive than its heaviest decay product, the charm quark. The quarks from a  $W$  decay will hadronize, the high mass of the parent  $W$  providing enough energy to its daughters so as to create a collimated jet of particles.

From an experimental viewpoint, the issue becomes finding an appropriate set of selection cuts to apply to the decay products to reconstruct a  $t\bar{t}$  pair. In general, jets from the decay will have their energies measured in the CDF calorimeters; by requiring a sufficiently high transverse energy,  $E_T$ , jets from

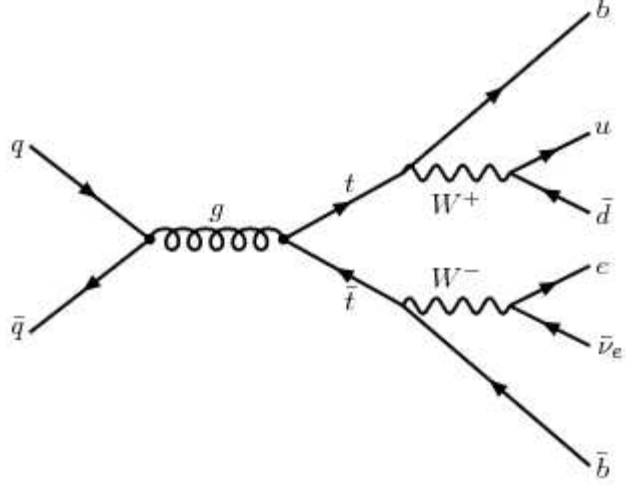


Figure 2.2: *Feynman diagram of  $t\bar{t} \rightarrow l + \text{jets}$*

the high-energy quarks of a  $t\bar{t}$  decay can be identified. Similarly, the charged lepton produced in the  $t\bar{t}$  decay can be identified by requiring it to have a sufficiently high  $E_T$ . As the  $\nu$  will escape the detector without interaction,  $t\bar{t}$  events in which at least one of the  $W$ 's decay leptonically can be identified by large missing transverse energy in the event,  $\cancel{E}_T$ . Finally, a  $b$ -tagging algorithm is used to identify jets originating from the  $b$ -hadrons of  $t\bar{t}$  decays [21]; this algorithm, along with the other selection cuts, is discussed in detail in Chap. 4.

In practice, top quark mass measurements are performed independently in three possible  $t\bar{t}$  decay channels, known as the “all-hadronic”, “lepton+jets” and “dilepton” channels. The all-hadronic channel consists of events selected for both  $W$ 's decaying to quarks. In the lepton + jets channel,  $t\bar{t}$  events are selected for one of the  $W$ 's decaying hadronically, and the other leptonically to a  $\nu$  and either a  $\mu$  or an  $e$ . Finally, in the dilepton channel, both  $W$ 's decay leptonically. A Feynman diagram of this decay can be seen in 2.2. Note that within the context of top mass analyses, a “leptonic decay” of a  $W$  excludes  $W \rightarrow \tau\nu$ , as the  $\tau$  will in turn decay making kinematic reconstruction of the  $t\bar{t}$  decay very difficult.

Different selection cuts, as well as different measurement issues, apply to the three channels. The advantage of looking for  $t\bar{t}$  events in which a  $W$  decays hadronically is higher statistics; with  $W \rightarrow \tau\nu$  decays treated as background, a  $W$  decays hadronically  $\simeq 3\times$  as often as it does leptonically. However, the leptonic decay of a  $W$  has a cleaner signature in the detector than the hadronic decay, as the background from non- $t\bar{t}$  multijet events is very high due to their large production cross section. For this reason, the all-hadronic channel tends to have good statistics and high backgrounds, the dilepton channel tends to have low statistics and low backgrounds, while the lepton + jets channel strikes a balance between the two<sup>1</sup>. In practice, this balance is optimal and causes the lepton + jets channel to yield the best measurements of the top quark mass. This can be seen in Fig. 2.3, in which the CDF measurements of the top quark mass as of March 2007, shortly before the approval of the measurement in this thesis, are shown.

### 2.3.3 Top Mass in Constrained Fits

The utility of a top mass measurement lies in the fact that, in conjunction with a measurement of the  $W$  mass, it can place a constraint on the mass of the Higgs boson through electroweak fits in the Standard Model. An interrelationship exists between the top, the  $W$  and the Higgs which underlies the fit, shown in the Feynman diagrams of Fig 2.4.

The  $W$  boson's coupling is affected by internal loops in its propagator. As its mass is defined by its coupling to the Higgs boson,  $gv/2$ , where  $v$  is the Higgs vacuum expectation value described in Sec. 2.1.3, these internal loops will therefore shift its mass. The high mass of the top quark means it has a large effect on radiative corrections to electroweak quantities; in particular, through

---

<sup>1</sup>However, at CDF the lepton + jets analyses have a lower background fraction than the dilepton analyses due to dilepton analyses not having a  $b$ -tag requirement for events

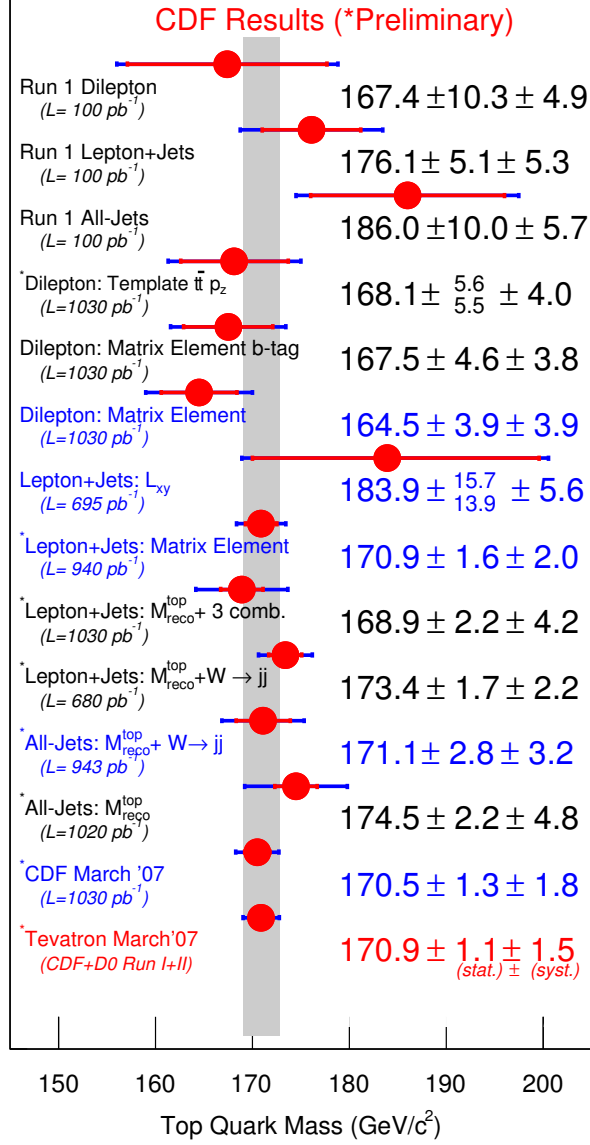


Figure 2.3: Measurements of the top quark mass at CDF as of March 2007. The CDF and the world averages are shown at the bottom.

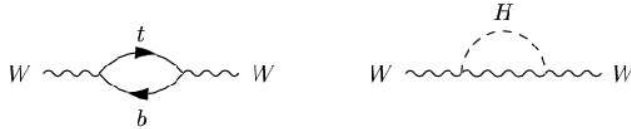


Figure 2.4: The  $W$  propagator, and hence its coupling to the Higgs' vacuum expectation value (i.e.,  $M_W$ ) is affected through internal loops with the top quark and the Higgs.

internal  $t\bar{b}$  loops, it shifts the  $W$  mass an amount proportional to  $M_t^2$  [22]. At the same time, through the emission and reabsorption of the Higgs in the  $W$  propagator, the mass of the  $W$  undergoes a shift proportional to  $\log M_H$ . Finally, the top quark has a strong Yukawa coupling to the Higgs, proportional to its mass and on the order of unity:  $\sqrt{2}M_t/v$ .

CDF and D0 have very strong programs both for the top mass measurement and the  $W$  mass measurement; better measurements of these two particles will provide a tighter constraint the Higgs mass in the future. In Fig. 2.5, one can see the central 68% confidence region for the world averages of the top and  $W$  mass measurements together, superimposed on diagonal contours of constant Standard Model Higgs mass. The contour of  $M_H = 114\text{GeV}/c^2$  lies just outside the 68% confidence region; however,  $M_H$  has been excluded experimentally by the CERN's Large Electron Positron (LEP) experiments below  $114\text{ GeV}/c^2$  to 95% confidence. Indeed, the most recent fit to high- $q^2$  global electroweak data predicts a Higgs lighter than that excluded by LEP. As can be seen in Fig. 2.6, the fit to the Higgs mass of  $76^{+33}_{-24}\text{ GeV}/c^2$  is just over  $1\sigma$  below the exclusion value. While this is certainly not statistical evidence of the inexistence of a Standard Model Higgs, it underlines the importance both of the effort at CDF and D0 to get better measurements of the top and  $W$  mass, as well as of exploring alternatives to the Standard Model.

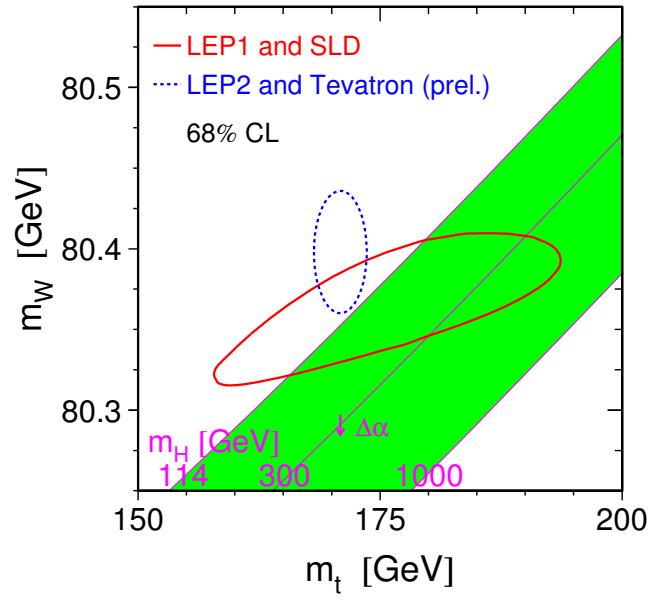


Figure 2.5: Central 68% confidence region of the world average top and  $W$  mass measurements, lying outside the contour of  $M_H = 114$  GeV/ $c^2$

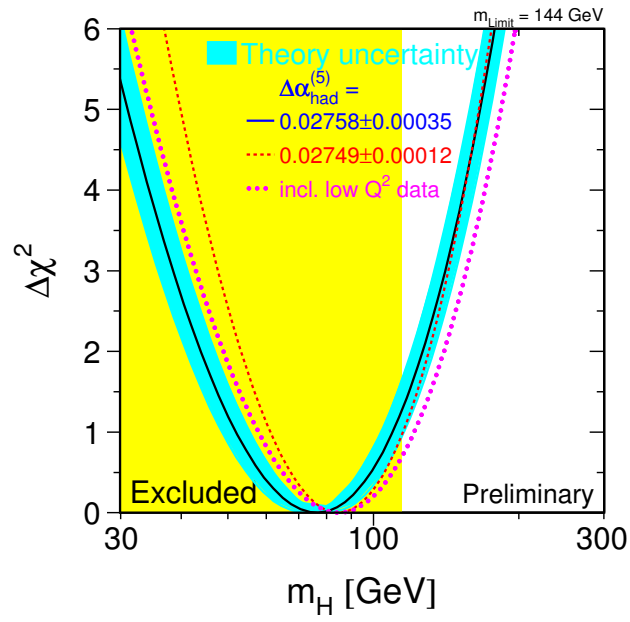


Figure 2.6:  $M_H$  vs.  $\chi^2$  of a Standard Model global fit to high- $q^2$  measurements; the blue region is LEP-excluded.





## **Chapter 3**

# **Experimental Apparatus**

## 3.1 Overview

The creation of a series of accelerators capable of achieving sufficiently high  $p\bar{p}$  collision energies to produce enough  $t\bar{t}$  pairs for a top mass measurement is a stunning technological achievement. Equally as impressive is the construction of a multipurpose detector which, among other things, is capable of allowing the kinematic reconstruction of the top quark's decay products so as to allow a measurement of the top quark's mass. In this chapter, the accelerator complex at Fermilab will be described, as well as the CDF detector, with a focus on the relevant subsystems in the detector used for the top mass measurement.

## 3.2 The Accelerator Complex

### 3.2.1 Overview of Luminosity

A large part of the CDF physics program concerns the search for, and measurements of, processes which have a very low cross section in comparison to the cross section of the inelastic  $p\bar{p}$  collisions. These processes include the production of  $t\bar{t}$  pairs, the production of a Standard Model (or non-Standard Model) Higgs, and various posited exotic phenomena. Aside from the problem of low cross section, often searches and measurements are severely hampered by the presence of background, necessitating tight selection cuts on candidate events which further increases the need for as many  $p\bar{p}$  collisions as possible. The measure of how many particle collisions are being produced per unit time at an accelerator is known as the “luminosity”. The mean number of occurrences of a given type of process produced at an accelerator,  $N_{events}$ , is found by integrating the luminosity over the time of the accelerator's beam delivery,

$$\mathcal{L}_{int} = \int \mathcal{L}(t) dt \tag{3.1}$$

and then multiplying this by the cross section of the process under consideration,  $\sigma$ :

$$N_{events} = \mathcal{L}_{int}\sigma \quad (3.2)$$

Luminosity can be expressed in a straightforward manner; given colliding bunches of protons and antiprotons in a circular synchrotron, it is

$$\mathcal{L} = \frac{f N_p N_{\bar{p}} N_b}{4\pi\sigma_x\sigma_y} \quad (3.3)$$

Here,  $f$  is the frequency of bunch collisions,  $N_p$  is the number of protons per bunch,  $N_{\bar{p}}$  is the number of antiprotons per bunch, and  $N_b$  is the number of bunches of each particle type (assumed to be the same, as is the case at the Tevatron). The term in the denominator,  $4\pi\sigma_x\sigma_y$ , is the area of the beam's (presumed) gaussian transverse profile; in practice a gaussian shape is usually a fairly accurate assumption. To maximize the number of collisions, one wants to focus the beams to have as low of a transverse profile as possible; techniques by which this is done at the Fermilab accelerator complex will be briefly touched on below.

### 3.2.2 Accelerating Protons before Tevatron Entry

A series of complex acceleration steps is needed to bring protons and antiprotons to the 980 GeV energies of collision; a sketch of the Fermilab accelerator complex is shown in Fig. 3.1. The protons to be used are taken from molecules of hydrogen gas; the molecules are split into hydrogen ions,  $H^+$  and  $H^-$ , through application of an electrostatic field. The  $H^-$  ions are accelerated by a Cockroft-Walton accelerator to 750 KeV, at which point further acceleration to 400 MeV is provided by a 150 m linear accelerator known as the “Linac”. Upon leaving the Linac, electrons are stripped from the  $H^-$  ions via a graphite foil, whereupon they enter a synchrotron whose diameter is roughly equal the

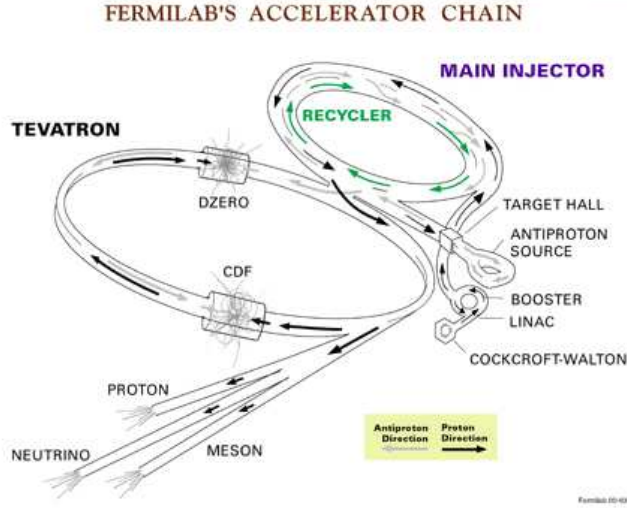


Figure 3.1: A sketch of the chain of accelerators used to produce  $\sqrt{s} = 1.96$  TeV  $p\bar{p}$  collisions.

length of the Linac, the “Booster”. The Booster accelerates protons to 8 GeV, after which they are sent to the Main Injector, a 3 km circumference ring which can accelerate protons to 150 GeV for injection into the Tevatron.

### 3.2.3 Creating and Accelerating Antiprotons

In order to create antiprotons, 120 GeV protons from the Main Injector are directed toward a nickel target. The production efficiency of  $\bar{p}$ ’s for every proton is  $\mathcal{O}(10^{-5})$ , and many other particles are created in the fixed-target collisions which must be distinguished from the  $\bar{p}$ ’s. A magnetized lithium lense is therefore used to take advantage of the fact that highly relativistic particles are bent differently in a magnetic field as a function of their mass,  $m$ , since  $p \simeq m\gamma c$  in this limit.  $\bar{p}$ ’s are produced at a rate of  $\sim 10^{11}/\text{hr}$ , and then sent to the Debuncher, which uses stochastic cooling to reduce the momentum spread of the antiprotons. The  $\bar{p}$ ’s are then stored in the Accumulator synchrotron before being added to the  $\bar{p}$  stash in the Recycler.

Sharing a tunnel with the Main Injector, the Recycler has proved to be

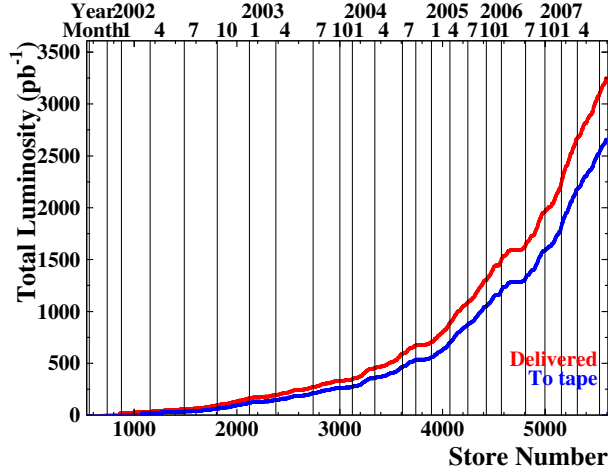


Figure 3.2: Total luminosity delivered by the Tevatron, as function of store #. Over  $3 \text{ fb}^{-1}$  has been delivered, with a goal through FY09 of  $8 \text{ fb}^{-1}$

invaluable in providing  $\bar{p}$ 's to the Tevatron in a manner which creates high luminosities. A relatively large number of  $\bar{p}$ 's can be used in a Tevatron collision run (known as a “store”) since the Recycler is capable of storing up to  $4\times$  as many antiprotons as the Accumulator. Entering use after a shutdown in late 2003, it has contributed to a tremendous growth in the luminosity delivered by the Tevatron, as can be seen in Fig. 3.2. In July 2005, electron cooling was used in the recycler for the first time. This technique, whereby electrons are injected into antiproton bunches in order to reduce the longitudinal momentum and position spread (or “emittance”) of the antiprotons, has allowed for better beam control and therefore higher luminosities. Finally, an interesting feature of the Recycler is its magnets; it is unique among accelerators of its size in that it uses permanent rather than superconducting magnets for beam control. The reason for this is that the opportunity cost of losing antiprotons due to a superconducting magnet quench was deemed too high.

### 3.2.4 The Main Injector and Recycler

The final acceleration stage before  $p$ 's and  $\bar{p}$ 's are injected into the Tevatron occurs in the Main Injector. Here  $p$ 's and  $\bar{p}$ 's are accelerated to 150 GeV, after which they are sent in bunches into the Tevatron. The Tevatron accelerates  $p$  and  $\bar{p}$  bunches from 150 to 980 GeV. A synchrotron, it is 6.3 km in circumference and uses 4.2 Tesla liquid-helium cooled superconducting magnets for beam control. Dipole magnets are used to steer the beam and quadrupole magnets are used to focus it to optimize luminosity. The beam consists of 36 bunches of  $p$ 's and  $\bar{p}$ 's each; bunches are designed to collide every 396 ns. The  $\bar{p}$  bunch contains approximately a trillion  $\bar{p}$ 's, while the amount of  $p$ 's in a bunch is roughly an order of magnitude larger. A store in the Tevatron typically lasts around a day, with the instantaneous luminosity continuously decreasing, first primarily through beam losses and then through  $p\bar{p}$  interactions.

On October 4, 2005, the Tevatron registered an instantaneous luminosity of  $140 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ , beating the record for instantaneous luminosity which had previously been set by the Intersecting Storage Rings (ISR) at CERN in 1982. As of August 8, 2007, the record has since more than doubled at the Tevatron, to  $291 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ , as the technologies underlying the accelerator complex have become better understood; a history of initial instantaneous luminosities can be seen in Fig. 3.3.

## 3.3 The CDF Detector

The CDF detector is a remarkable experimental apparatus. Possessing a roughly cylindrical geometry consisting of concentric subsystems around the Tevatron beampipe designed for different aspects of  $p\bar{p}$  collision event reconstruction, it is approximately 12 m long, with a similar diameter, and weighs approximately 5000 tons. A cross-sectional sketch of the detector is shown in

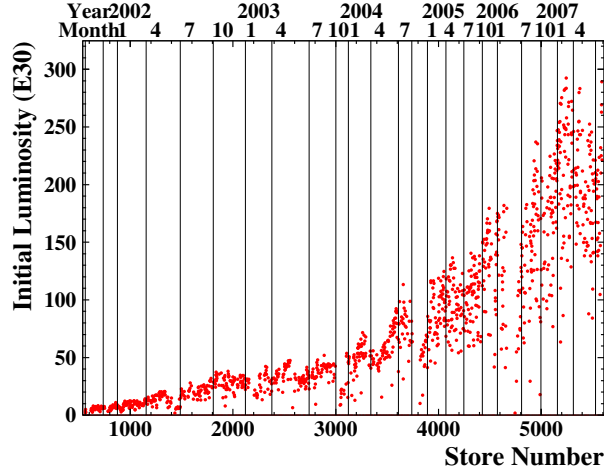


Figure 3.3: *Initial instantaneous luminosities, by store, during Tevatron's Run II.*

Fig. 3.4; a person is included in the drawing to give a sense of scale. The CDF detector is multipurpose, and several hundred physicists from all over the world use the detector for a variety of physics, including top and  $W$  mass measurements, exotics searches, flavor and QCD physics. Not all analyses utilize all subsystems, and the following description of the detector will focus on those subsystems used to reconstruct  $t\bar{t}$  events: the tracking system, used for lepton momentum measurements and  $b$ -tagging, the calorimetry, used for particle jet and electron energy measurements, and the muon chambers, used to track muons which typically punch through the calorimetry unlike other particles. Reconstruction details will be covered in depth in Chapter 4.

As the rate of bunch interaction events is 1.7 MHz at the Tevatron, it is physically impossible to use the detector to reconstruct and store them all. However, the vast majority of these interactions are uninteresting physically. For these reasons, CDF possesses a three-tiered trigger system designed to identify events of possible physics interest, saving to tape about 75 events out of the 1.7 million produced every second. In addition to the detector subsystems of interest, this section will address the design of the trigger system.



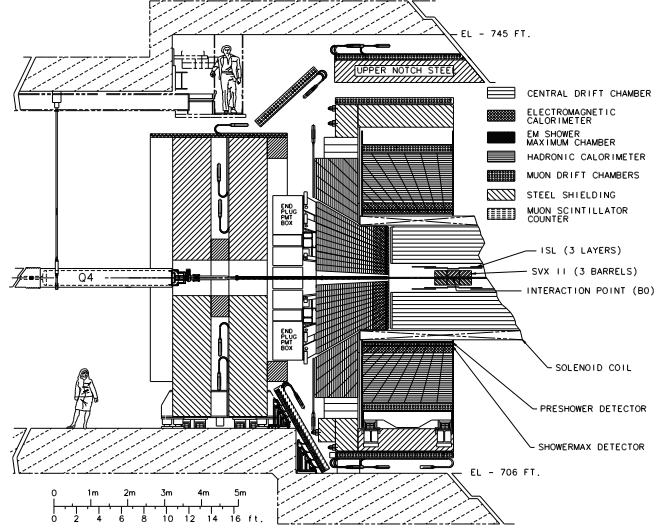


Figure 3.4: *Cross-sectional sketch of the CDF detector, from the CDF Run II Technical Design Report*

### 3.3.1 Geometry

When describing the geometry of the CDF detector, cartesian, cylindrical and polar coordinates are used interchangeably. The correspondence between these coordinates and the detector's real-world orientation is as follows:

- $+z$  is along the proton beam direction, pointing east
- $+y$  points away from the earth
- $+x$  points north
- “ $r$ ” is the radius from the beamline,  $r = \sqrt{x^2 + y^2}$
- $\phi = \tan^{-1}(y/x)$
- $\theta = \cos^{-1}(z/(\sqrt{x^2 + y^2 + z^2}))$
- $\eta = -\log(\tan(\frac{\theta}{2}))$

In practice,  $\eta$  – the “pseudorapidity” – is used more often than  $\theta$ , in this thesis specifically and in high energy physics generally. The reason is the following: the variable “rapidity”, defined as  $Y = \frac{1}{2} \log((E + p_z)/(E - p_z))$ , where  $E$  is the energy of a particle and  $p_z$  its z-momentum, has an approximately even distribution for particles exiting the interaction region. Its geometrically-defined analog is  $\eta$ , which has a similar value to the rapidity for a particle in the relativistic limit.

### 3.3.2 Tracking

The CDF tracking system, concentric about the beampipe, consists of a silicon microstrip system (divided into three subdetectors, the L00, the SVXII and the ISL) and the COT (Central Outer Tracker), an open-cell drift chamber. Both are immersed in a 1.4 Tesla solenoidal field allowing for position and hence momentum resolution of charged particles.

#### Silicon Microstrip Detectors (L00, SVXII and ISL)

The silicon microstrip detectors [23], a cross sectional view of which is shown in fig. 3.5, are used for very high precision tracking of charged particles. They are designed to be radiation tolerant due to the high particle fluences just outside the beampipe, and are critical for use in the  $b$ -tagging algorithm used in the selection cuts of this analysis. The algorithm uses individual particle tracks to extrapolate interaction vertices and is described in Section 4.5. Additionally, they can provide tracking for particles which are too far forward to be detected by the COT.

Located immediately outside the beampipe (at  $r \sim 1.3$  cm) is the L00 detector, consisting of single sided silicon wafers with axial silicon microstrips. Beyond this, extending from  $z = -45$  to  $+45$  cm and  $2.5 < r < 10.6$  cm, and covering  $|\eta| < 2$ , is the SVXII detector. The SVXII detector is built from three

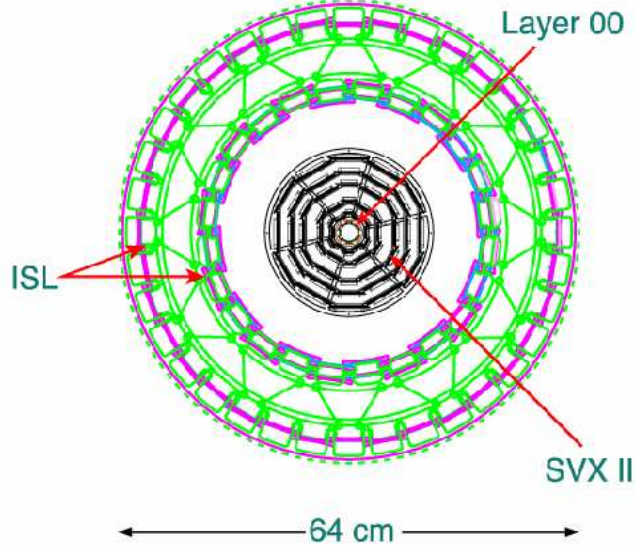


Figure 3.5: *Cross-sectional view of silicon tracking system (L00, SVXII and ISL)*

closely spaced identical “barrels”. Each barrel is composed of five layers, and each layer is composed of double sided silicon sensors, arranged into 12 wedges in  $\phi$ . Each sensor has a side consisting of strips parallel to the beam axis – “axial strips” – and a side consisting of strips either at a  $1.2$  or  $90^\circ$  angle relative to the parallel strips – “stereo strips”. The stereo strips provide some measure of position resolution along the  $z$ -axis, though this is on the order of hundreds of  $\mu\text{m}$ , over an order of magnitude larger than the subsystem’s  $r$ - $\phi$  resolution.

Outside of the SVXII detector is the ISL, or Intermediate Silicon Layer. The ISL consists of three layers, a central layer at  $r = 22$  cm and  $|\eta| < 1$ , and two forward layers, at  $r = 20$  and  $28$  cm and  $1 < |\eta| < 2$ . Each layer consists of double sided silicon sensors, with a set of axial strips and a set of small-angle stereo strips. The ISL provides continuity between the tracking capabilities of the SVXII and the COT.

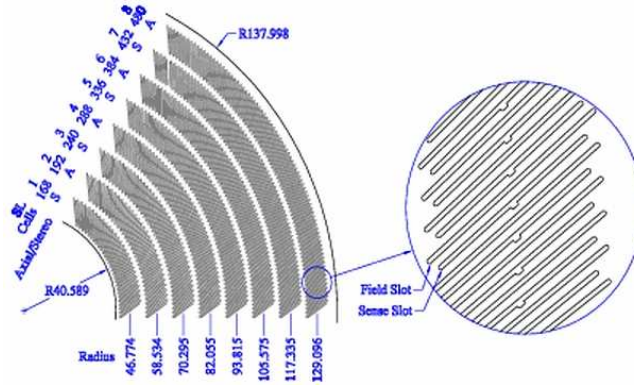


Figure 3.6: *Partial cross-section of the COT. Every superlayer is 12 sense wires deep, and is segmented into chambers, the number of which at each superlayer is written on the picture.*

## COT

The COT [24] is an open-cell drift chamber capable of tracking charged particles at  $|\eta| < 1$ ; extending from  $r = 40$  to  $r = 137$  cm, it is the Run II replacement of CDF’s Run I drift chamber, the CTC (Central Tracking Chamber). It consists of 8 “superlayers”, shown in Fig. 3.6, where each superlayer is 12 sense wires deep. Alternating layers either have the wires placed axially or at a small-angle stereo ( $+ \text{ or } -3^\circ$ ); one of the goals for the COT was that it would provide better  $z$  resolution than the CTC. As charged particles pass through the COT, they ionize a gas composed of an argon-ethane- $\text{CF}_4$  combination creating avalanches of electrons which drift to the sense wires. The drift time is designed to be no more than 100 ns, as the original plan had been to operate the Tevatron with 132 ns spaced bunch crossings for CDF Run II [25].

### 3.3.3 Solenoid

The entire tracking system is contained within a superconducting solenoid. Cooled by liquid helium, the solenoid is made of an aluminum supported niobium-titanium alloy which can operate with a current of 5000 A through its 1164

windings to produce a magnetic field of 1.5 T. In practice, however, it is operated at 4650 A for a field of 1.41 T, which points in the -z direction. It is located outside the tracking system with an inner radius of 1.4 m, and is approximately 25 cm thick.

### 3.3.4 Calorimetry

While the purpose of the tracking system is to measure the location of individual charged particles and thus their momenta, the calorimetry at CDF aims to measure the energies of particles. Particles entering the calorimeters have their energies measured through the shower of decay products they produce through interacting with the material of the calorimeter. While the tracking system, then, is designed to be made of as little material as possible so as not to alter the path of a charged particle through interaction, the calorimeters are designed to fully capture the energy of the particles which enter them. This is done through use of scintillating material interspersed with a heavy absorbing material.

The calorimeters are located outside of the solenoid and possess a nearly  $4\pi$  solid angle coverage, covering almost all of  $\phi$  and  $|\eta| < 3.6$ . There are two different calorimeter types used; right outside of the solenoid is the electromagnetic calorimetry [26], designed to measure the energies of photons and electrons. The electromagnetic calorimeter is designed to be many radiation lengths,  $X_0$ , deep - about  $19 X_0$  in the central region of the detector, and slightly more in the forward region; further statistics can be found in Table 3.1.  $X_0$  is the amount of material through which electrons must pass to lose on average all but  $1/e$  of their energy through bremsstrahlung, or  $7/9$  the mean free path before a photon undergoes pair production; it has dimensions of mass/area [12]. As all charged particles aside from the electron are typically too massive to lose much energy through the bremsstrahlung the electromagnetic calorimeter is designed to

Calorimeter	Resolution (%)	Coverage ( $ \eta $ )	Depth
CEM	$13.7/\sqrt{E_T} \oplus 2$	0-1.1	$19X_0$
PEM	$16/\sqrt{E_T} \oplus 1$	1.1-3.6	$21X_0$
CHA	$50/\sqrt{E_T} \oplus 3$	0-0.9	$4.5\lambda$ (+WHA)
WHA	$75/\sqrt{E_T} \oplus 4$	0.8-1.2	$4.5\lambda$ (+CHA)
PHA	$80/\sqrt{E_T} \oplus 5$	1.2-3.6	$7\lambda$

Table 3.1: *Statistics for the calorimeters.*

induce, and neutral particles don't lose energy through bremsstrahlung, their energies have to be measured by the hadron calorimetry [27]. The hadron calorimetry in the central region is 4.5 interaction lengths,  $\lambda$ , deep, where  $\lambda$  is the mean distance a hadron must travel before undergoing a strong interaction with a nucleus; in the plug region, this depth increases to 7  $\lambda$ . These systems are designed to measure the energies of charged and neutral hadrons through strong interactions with the nuclei of their material. Typically the only known particles aside from neutrinos which can consistently escape the calorimeters are muons; how this is dealt with will be described in Sec. 3.3.5.

Overall, the calorimeter system consists of a central electromagnetic and hadronic calorimeter (CEM and CHA), plug electromagnetic and hadronic calorimeters (PEM and PHA), and finally, wall hadronic calorimeters (WHA) which fill in a region in  $|\eta|$  that the CHA and PHA do not, as can be seen in Fig. 3.4. The CEM consists of 478 towers, segmented by  $15^\circ$  in  $\phi$  and with 10 on either side of the interaction region up to  $|\eta| < 1.1$ ; the reason there are 478 rather than 480 towers is due to the fact that in two locations where towers would have been, space has been made for components of the solenoid's cryogenic system. The CEM is 31 layers of lead and scintillator deep. Energies are read out from the scintillators as wavelength-shifting fibers send the scintillation light to two photomultiplier tubes per tower.

Outside the CEM is the CHA, best envisioned in conjunction with the WHA with which it shares towers. Some of the towers of the CHA and WHA match up with the towers in the CEM; the segmentation is  $15^\circ$  in  $\phi$ , and the  $|\eta|$  coverage extends to 1.2 with 12 total towers, 6 of which are wholly part of the CHA, 3 of which are wholly part of the WHA, and 3 of which are shared. The CHA is constructed of 32 layers of steel and scintillator, and the WHA of 15 layers of steel and scintillator.

While the CEM extends to  $|\eta| = 1.1$  and the WHA extends to  $|\eta| = 1.2$ , the plug calorimeters continue the CDF detector's calorimetric coverage to  $|\eta| = 3.6$ ; the PEM lies in the region  $1.1 < |\eta| < 3.6$ , and the PHA lies in  $1.2 < |\eta| < 3.6$ . Both calorimeters are made of the same scintillator/absorber materials as the more central calorimeters. However, whereas the CEM, CHA and WHA were all constructed for Run I, the plug calorimeters were constructed for Run II, as the Run I calorimeters in the same  $|\eta|$  region were too slow for the originally planned 132 ns bunch crossing time. The PEM is 23 layers deep, and consists of 12 towers in  $\eta$ . The four tower groups which are in the most forward region are segmented into  $15^\circ$  much as those towers in the CEM, CHA and WHA; however, the other eight tower groups have a finer  $\phi$  segmentation of  $7.5^\circ$ .

Finally, it should be noted that within the physical region of the calorimeters are subsystems which provide valuable service to event reconstruction. In both the CEM and PEM, around  $6X_0$  into the detector from the beampipe are “showermax” detectors, consisting of wires and scintillator strips (in the case of the CEM's showermax) or scintillator strips alone (in the case of the PEM's showermax). The name “showermax” derives from the fact that a particle's resulting shower in the calorimeter reaches its peak at about this depth. The detectors are used to provide 2-dimensional information on the showers in the calorimeters.

### 3.3.5 Muon System

Outside of the calorimeters is the muon system, consisting of thousands of single-wire drift tubes run in proportional mode, as well as thousands of scintillator strips, with steel shielding interspersed to absorb particles which could fake a muon track. Whereas electrons and photons are typically stopped in the EM calorimetry, and hadrons are stopped due to the several interaction lengths of material in the hadronic system (less than 1% of pions make it through), muons of more than 1-2 GeV (and certainly muons of interest in  $t\bar{t}$  decays) tend to “punch through” the calorimeter system. The reason for this is that they are minimum ionizing particles from a few hundred MeV to several TeV, and also lose far less energy through bremsstrahlung than the much lighter electrons. In an event, when tracks are identified in the single-wire tubes of the muon system, they are typically matched to hits seen in the tracking system for muon reconstruction.

The muon subsystems include the central muon detector (CMU), the central muon upgrade detector (CMP), the central muon extension detector (CMX), as well as the intermediate muon detector (ISU). Only the first three subsystems will be discussed, as the ISU is not used for muons in this analysis. In all three systems, the drift tubes are stacked four deep. The coverage in  $\eta$ - $\phi$  space of these muon systems is shown in Fig. 3.7.

#### CMU

The central muon detector, CMU, is located outside the CHA, separated from it by steel shielding. The CMU consists of 2880 single wire proportional chambers; each chamber is  $6.35 \times 2.68 \times 226$  cm in size. The CMU is  $\phi$ -symmetric and provides full coverage up to  $|\eta| < 0.6$ .



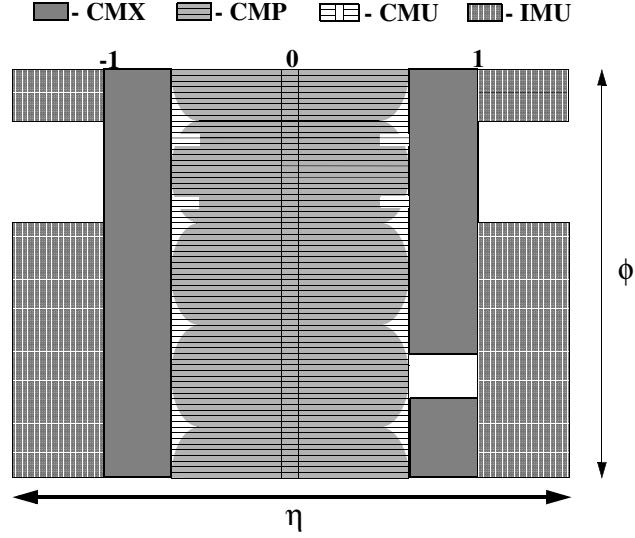


Figure 3.7: Coverage, in  $\eta$ - $\phi$  space, of the muon system.

### CMP

The CMP similarly provides coverage up to  $|\eta| < 0.6$ , and is located beyond an additional 60 cm of steel provided by the solenoid return yoke from the CMU. It is composed of 1076 single wire proportional chambers of  $2.5 \times 15 \times 640$  cm. Unlike the CMU and the subsystems described so far in this thesis, it is not  $\phi$ -symmetric, but rather has a boxed shape due to the fact that it is built around the box-shaped return yoke. Due to the unique geometry of the CMP, the arrangement of the drift tubes is not symmetric about the beam, and coverage is not complete in  $\eta$ - $\phi$  space, as seen in Fig. 3.7. Still, covering the same general  $\eta$  region as the CMU, it offers enhanced muon identification capability not simply through its additional drift chambers but also due to extra shielding between it and the CMU.

### CMX

Extending coverage of the muons is the CMX, which covers  $0.6 < |\eta| < 1.0$ . 2208 drift tubes of dimensions  $2.5 \times 15 \times 180$  cm are used for the east and west

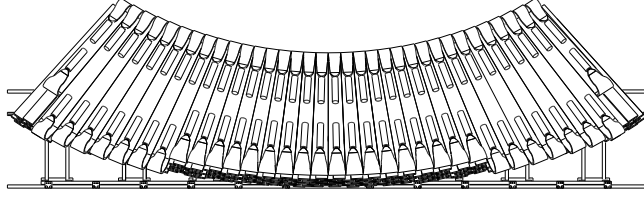


Figure 3.8: *Lower chambers for the Central Muon Extension and Scintillators, known as the “miniskirt” and possessing a planar fan geometry.*

portions of the CMX. The shape of each half is that of a truncated cone, with the exception of the detector quadrant closest to the floor; here the constraint provided by the floor yields a different geometry in the form of the fan-shaped “miniskirt” region (see Fig. 3.8). Each half of the CMX is identical, with the exception of the  $30^\circ$  in  $\phi$  highest off the ground in the east portion, which is left empty so as to provide room for cryogenic utilities. The west portion had also been empty during Run I but had tubes installed for Run II; this region is referred to as the “keystone”.

No description of the muon system is complete without mention of the scintillators. On the outsides of the CMP and the CMX are scintillator tiles; 269 tiles are used in the CMP, 272 in the CMX. These tiles are in place due to the fact that the drift time in the muon system drift tubes can be as high as  $1.4 \mu\text{s}$ , between the length of 3-4 396 ns bunch collisions. While the occupancy of the drift tubes is low enough that there’s no danger of event pileup, the much faster timing of the scintillators is needed to allow the association of muon stubs with the proper event.

### 3.4 Trigger System

Due to bandwidth limitations, only a few events out of every hundred thousand are stored to tape. The decision as to which events are to be recorded is made by a three-tiered trigger system, composed of the Level 1 and Level

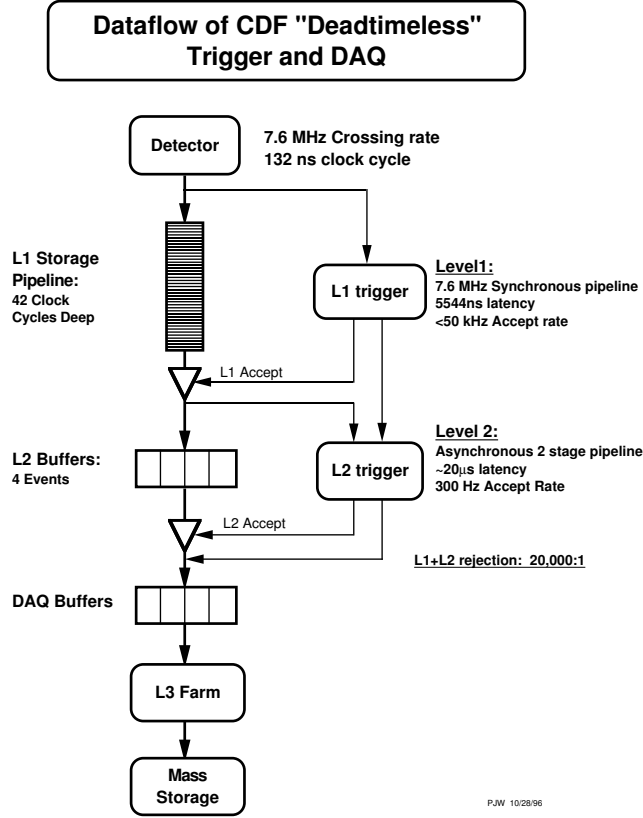


Figure 3.9: *Dataflow of the CDF trigger system. The acceptance rate is drastically lower than the bunch collision rate.*

2 hardware triggers and the Level 3 software trigger. The system's layout is shown in Fig. 3.9.

### 3.4.1 Level 1 Trigger

The Level 1 trigger, designed to handle the initially planned Run II bunch crossing rate of 132 ns, stores all events in a 42-event deep synchronous pipeline. Its pass rate is around 20-25 KHz; taking only around  $4 \mu\text{s}$  to make a decision, it is completely deadtimeless. There are three primary components to the L1 trigger: the “XFT” (eXtremely Fast Tracker), which records the pattern of charged particle hits in the COT; raw energy measurements in individual calorimeter towers; and the matching of hits in the muon systems to those recorded by the

XFT. There are numerous L1 triggers, any of which, if satisfied, will send an event to the L2 trigger. Each of these L1 triggers can be prescaled independently in the event of very high instantaneous luminosity.

### **3.4.2 Level 2 Trigger**

The Level 2 trigger is capable of more sophisticated event reconstruction; different features of the L2 event are examined by a set of Alpha processors which allow for a passing rate of a few hundred Hz. Events to be handled by the L2 trigger are stored in one of four asynchronous buffers; as an event is held in a buffer until it is reconstructed and examined, this trigger is not completely deadtimeless, although its deadtimeless rate is only  $\sim 10\%$ . At L2, calorimeter energy clusters are measured through the summing of individual tower energies; additionally, further positional information on the calorimeter clusters is obtained through use of the showermax information. Tracking is performed through use of the “SVT” (Silicon Vertex Tracker) algorithm, which reconstructs tracks through using silicon hits in the neighborhood of XFT tracks. L1 muon trigger information is also employed in event reconstruction.

### **3.4.3 Level 3 Trigger**

The Level 3 trigger consists of C++-implemented software running on a set of Linux PCs. It performs further event reconstruction and selection, providing a passing rate of  $\sim 100$  Hz. Events which are passed are saved to tape in the Feynman Computing Center at Fermilab, whereupon they undergo further formatting in order to be usable to the various CDF analyses.



## Chapter 4

# Event Reconstruction and Selection

## 4.1 Overview

As mentioned in Sec. 2.3,  $t\bar{t}$  events have a low cross section at the Tevatron; only a few can be expected to be produced in a 24-hr store. Fortunately, the leptonic  $W$  decay in the  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  decay channel provides a striking signature in the form of a high- $p_T$  charged lepton. The three triggers used to obtain the dataset for this analysis take advantage of this; they are referred to as the “CEM” trigger, corresponding to an electron in the CEM, the “CMUP” trigger, for a muon in the CMU and CMP, and the “CMX” trigger for a muon in the CMX. Events which pass one of these triggers are subjected to a further series of offline selection cuts. The primary cuts require the following:

- A high- $E_T$  electron or high- $p_T$  muon, with a stricter and more sophisticated definition of what constitutes an acceptable electron or muon than that found in the trigger.
- Four and only four high- $E_T$  jets in the  $|\eta| < 2$  region of the calorimeters, corresponding to the hadronization of the four quarks; these will hence be referred to as “tight” jets.
- A large imbalance in the transverse energy of the event,  $\cancel{E}_T$ , due to the undetected high- $p_T$  neutrino.
- At least one of the four jets must be tagged as coming from a  $b$ -quark.

How we use the detector to reconstruct these quantities we select on will be discussed in this section; additionally, selection cuts used for further identification capability will be described. The full set of requirements for  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  candidate events are simply those of the CDF Top Group [28], excepting for the fact that the Top Group selection cuts allow for  $> 4$  tight jets.

## 4.2 Electron Reconstruction

In events with the CEM trigger, at L1 XFT-reconstructed electron tracks in the COT are required to point to calorimeter towers with significant amounts of energy; this requirement is refined at L2 through the replacement of calorimeter tower energies with calorimeter cluster energies<sup>1</sup>. At L3, the electron is required to have a fully-reconstructed transverse energy of 18 GeV, with a matching COT track of  $p_T > 9$  GeV/c; the showermax plays a role in this matching. The ratio of the cluster energy measured in the hadronic calorimeter to that in the electromagnetic,  $E_{had}/E_{em}$ , must be less than 0.125.

Offline, the basic requirement is that the measured cluster have  $E_T \leq 20$  GeV. There are a number of important subsidiary requirements, however. The  $E_{had}/E_{em}$  requirement is tightened to  $0.055 + 0.00045E$ ; the energy dependence here accounts for the fact that higher-energy electromagnetic showers created by electrons can leak more energy into the hadronic calorimeter than lower energy showers. The ratio of cluster energy to matched track momentum,  $E_{em}/p_{track}$ , as a measure of the isolation of the track, must be less than 2 if  $E_T < 50$  GeV; furthermore,  $p_{track}$  must be  $> 10$  GeV/c. As a measure of the isolation of the shower, the ratio of energy in towers adjacent to the tower into which the charged lepton passes to the energy in that tower itself must be  $< 0.2$ . Finally,  $\gamma \rightarrow e^+e^-$  conversions are rejected by searching for an antiparticle with the appropriate kinematics. If a third electron or positron is found, this is considered indicative of pair production from a bremsstrahlung photon radiated by the initial charged lepton, and the event is not rejected.

---

<sup>1</sup>see Sec. 4.4 for more on clustering



### 4.3 Muon Reconstruction

In events with CMUP or CMX triggers, the challenge is to identify and measure high- $p_T$  muons. These two triggers have much in common; at L1 and L2, tracks left by muon candidates in the muon system, known as “stubs”, must be matched to tracks left in the COT. In the case of the CMUP trigger, a stub in the CMU must be corroborated by a stub in the CMP; this selects against charged pions which may make it into the CMU. By L3, along with the requirement of a match, the muon must have a reconstructed  $p_T > 18$  GeV/c.

Offline, the primary requirement on the muon is that it have a  $p_T > 20$  GeV/c. Additionally, the energy deposited in the calorimeter by the candidate muon must be consistent with that of a minimum ionizing particle. As a means of eliminating secondary muons, the reconstructed track of the muon candidate must have an impact parameter less than 0.2 cm if only the COT was used for the track, or 0.02 cm if the silicon detector, with its much higher precision tracking, was used. Cosmic-ray muons are vetoed by comparing the high-precision timing afforded by the scintillator strips to the bunch crossing times. Finally, to ensure muon isolation, a requirement that the ratio of  $E_T$  deposited in a cone of  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$  around the muon track to the muon’s  $p_T$  be less than 0.1 .

### 4.4 Jet Reconstruction

As described in Sec. 2.1.2, colored quarks and gluons interact through the strong force so as to always be arranged in color neutral configurations known as “hadrons”; a consequence of this is that the four quarks from the  $t\bar{t}$  decay must be reconstructed in the detector in the form of particle jets, consisting of the hadrons and their decay products resulting from the fragmentation and

hadronization of the initial quarks<sup>2</sup>. Thus, unlike the case with the charged leptons in  $t\bar{t}$  events, it is not possible to directly measure the kinematics of the partons (quarks or gluons). However, reconstruction of this quantity can occur by measuring the raw energies in the calorimeter towers through which a jet is thought to pass, and then applying a set of corrections to those energies to recover the original parton kinematics [36]. Unfortunately, the systematic uncertainties associated with these corrections result in an overall systematic uncertainty of about 3% for the characteristic jet energies seen in  $t\bar{t}$  decays, which can have a profound effect on the measurement of the top quark mass; see Fig. 4.1 for details. How we handle this systematic in the analysis is described in Chap. 5.

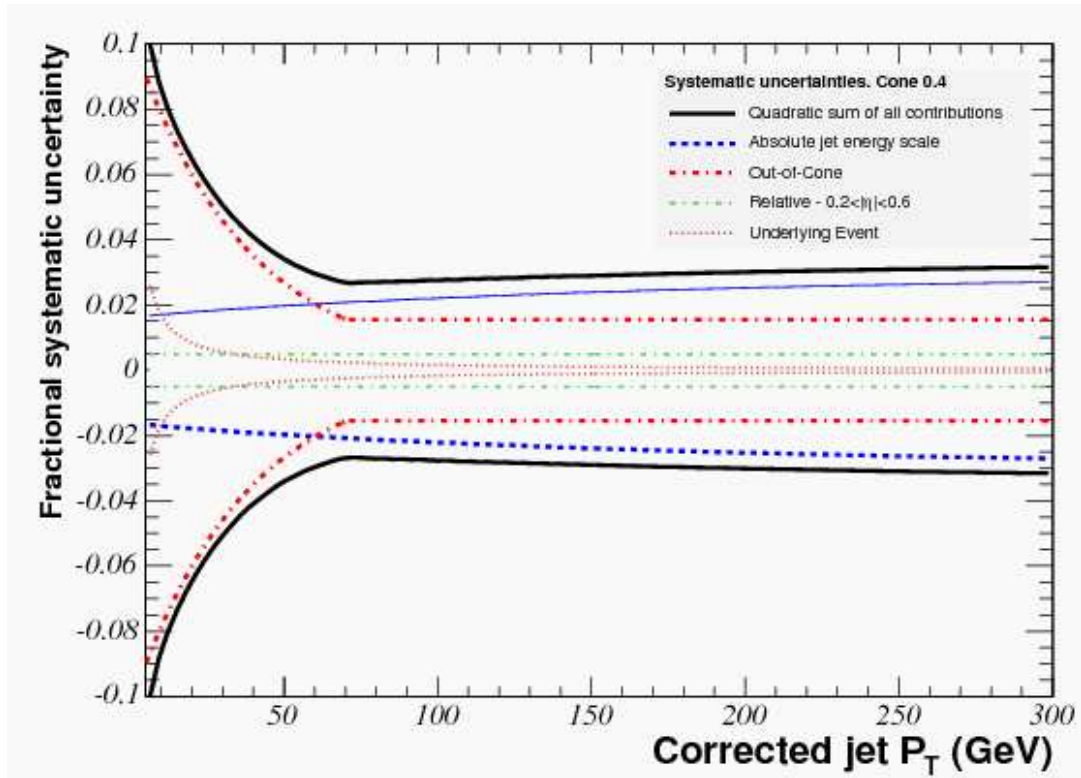


Figure 4.1: *Systematic errors on jet energy corrections as function of jet  $p_T$ .*

<sup>2</sup>The exception to this rule is, of course, the top quark, which is massive enough to decay weakly before it has a chance to hadronize

#### 4.4.1 Jet Clustering and the Raw Measurement

In order to identify a particle jet in the calorimeter system, an algorithm designed to locate the towers containing a jet’s energy called “JETCLU”, short for “jet clustering” [29] is employed. The algorithm begins with the identification of a “seed tower”, which is a tower with a measured raw  $E_T > 1$  GeV. The reason  $E_T$  rather than  $E$  is used is related to the manner in which corrections are applied to the energy, described below; we have  $E_T = E \sin \theta$ , where  $\theta$  is the polar angle obtained by the calculated production vertex of the jet, the center of the tower in question, and the z-axis.

Once the seed tower is identified, all towers whose centers are located within a radius of  $\Delta R < 0.4$  and which have  $E_T > 0.1$  GeV are provisionally taken to contain some of the energy of the jet; the centroid of the jet is then calculated as an average in  $\eta$  and  $\phi$  of the towers, weighted by each tower’s  $E_T$ . The towers which satisfy the above requirements relative to the *new* centroid are exposed to the same algorithm, and the process iterates until a constant set of towers is found, whereupon their energies are summed. Separate jets which share a sufficient amount of energy between the same towers are merged. With the raw energy in a cluster measured for a jet, a series of corrections then need to be applied.

#### 4.4.2 The Relative Correction

The first correction is designed to account for differences in jet response based on the detector region; as one would expect, jets passing through regions where there are cracks due to two calorimeters meeting yield somewhat different energy measurements than jets in the fiducial regions of the central calorimeters. This correction is applied by taking dijet events in which one jet, the “trigger”, is required to be in a fiducial region of the central calorimeter, defined as  $0.2 < |\eta| < 0.6$ . The other jet may be in any  $\eta$  region covered by the calorimetry.

Due to the fact that the vector sum of the two jets  $E_T$ 's should be 0, it is possible to use this approach to find a correction such that jets outside of the  $0.2 < |\eta| < 0.6$  region have their  $E_T$ 's scaled to correspond to jets in that region.

#### 4.4.3 Multiple Interactions Correction

The next correction is designed to account for the effect of energy from other  $p\bar{p}$  interactions in the same bunch crossing as the event under study being added to a jet cluster. This correction is obtained via measurements of energy in random cones of  $\Delta R < 0.4$  in minimum bias events with different numbers of reconstructed vertices. Given the instantaneous luminosities achieved by the Tevatron during the period in which the dataset for this analysis was collected compared to current luminosities, this correction has an effect of only a couple of tenths of a GeV.

#### 4.4.4 Absolute Correction

In the central calorimeter, a nonlinear response in the measured jet energy to the true underlying parton energy is expected; for this reason, an “absolute correction” is applied to jet energies which have already undergone the previous two corrections. This correction is determined by taking dijet events in Pythia Monte Carlo (MC) and fitting to distributions of the difference between measured and true jet  $p_T$  from these events. In this case, systematic uncertainties arise due to differences in the  $p_T$  distributions of jet particles for Pythia vs. data which are incorporated into the overall jet energy systematics.

#### 4.4.5 $E_T$ Selection Requirement

There are two other jet corrections, from the energy deposited inside the cluster by the underlying event and the energy from the jet deposited outside

the cluster. As these corrections are calculated primarily from events involving purely QCD processes, they cannot be unreservedly used for jets from the  $t\bar{t}$  decay; for this reason we employ transfer functions which will bring us from the true  $p_T$  of the parton to the measured jet momenta; this is described in Sec. 5.6.

Although the jets used as inputs to the analysis use the relative, multiple interaction and absolute corrections, when selecting candidate  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  events, the transverse energy of the jets is only corrected using the relative and multiple interaction corrections. By this definition, the  $E_T$  of a tight jet must exceed 15 GeV. Jets which have been corrected using the relative and multiple interaction corrections are said to have been corrected to “Level 4” or “L4”; jets which have the absolute correction as well have been corrected to “Level 5” or “L5”.

#### 4.4.6 Missing $E_T$ ( $\cancel{E}_T$ )

As the transverse energies of the incoming proton and antiproton which collide and create  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  events are very low relative to their total energy of 980 GeV, an imbalance in transverse energy will be found when the neutrino in the decay escapes undetected in this type of event. In selecting events with a large missing  $E_T$  ( $\cancel{E}_T > 20$  GeV) we seek this effect. The basic formula for  $\cancel{E}_T$  is as follows:

$$\cancel{E}_T = \left| \sum_{i=1}^{i=\#of towers} E_T \hat{n}_T \right| \quad (4.1)$$

i.e., simply the magnitude of the vector sum of the raw  $E_T$ ’s measured in every calorimeter tower; here  $\hat{n}_T$  is the unit vector in the x-y plane pointing from the origin to the tower. However, two additional corrections to this quantity are made. One is to account for muons, which, unlike other particles, typically deposit only a small fraction of their energy in the towers, and the other is to add the additional information gained from jets which have had the corrections described in the previous section applied to their raw energies.

## 4.5 Secondary Vertex $b$ -tagging

Many non- $t\bar{t}$  events which contain a  $W$  and four jets will be able to pass the aforementioned selection cuts. However, a relatively low proportion of them will contain two  $b$  quarks in their final state like the  $t\bar{t}$  decays. Therefore, a way to tag a jet as originating from a  $b$  quark would be very useful in selecting  $t\bar{t}$  candidate events.

Fortunately, it is possible to take advantage of the tracking capabilities of the silicon system to do just that [30]. The  $b$ 's from top decay will convert into B-hadrons which have energies on the order of several 10's of GeV, due to the high mass of the parent top. This high energy slows the decay time of the B, which is  $\sim 1.5$  ps in the B's rest frame, by a significant relativistic time dilation factor allowing the B to typically travel millimeters before decaying into a jet of particles. By taking a jet in a  $t\bar{t}$  candidate event and attempting to reconstruct the vertex of origin (the "secondary vertex") of the jet through its charged particles, it is possible to determine if the jet originated a sufficient distance from the primary vertex of the interaction to be tagged as a  $b$  jet<sup>3</sup>.

The algorithm can be outlined as follows: first, the charged particle tracks in the jet under study are subjected to selection cuts to insure that a quality secondary vertex can be reconstructed. There must either be at least 3 tracks with  $p_T \geq 0.5$  GeV/ $c$  where at least one of the tracks must be  $> 1$  GeV, or there must be at least two tracks with  $p_T \geq 1$  GeV/ $c$ . Once the secondary vertex is calculated using the tracks, the distance in the x-y plane between the primary and secondary vertices is projected onto the direction of the jet; this quantity is referred to as  $L_{2D}$  (see Fig. 4.2 for more details). A jet is tagged if  $L_{2D} > 7.5\sigma_{L_{2D}}$ , where  $\sigma_{L_{2D}}$ , the error on  $L_{2D}$ , is around 190  $\mu\text{m}$ ; additionally, some  $\chi^2$  requirements related to the track fitting must also be satisfied.

---

<sup>3</sup>or a jet originating from a charm quark,  $c$ ; collectively  $b$  and  $c$  are referred to as "heavy flavor" quarks

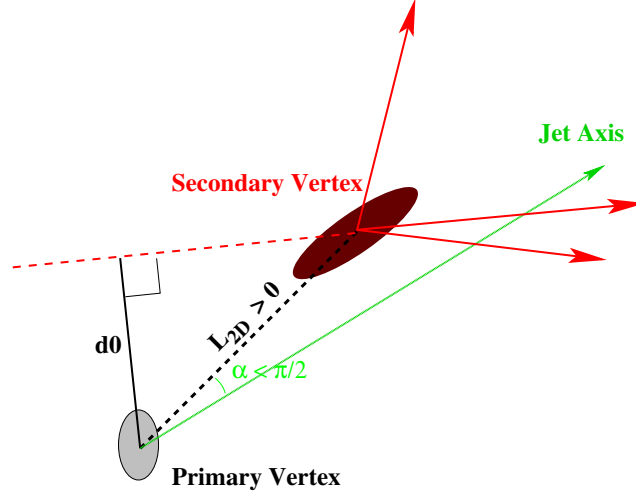


Figure 4.2: A view of a  $b$  jet in the  $x$ - $y$  plane. The primary means of determining whether a jet originates from a heavy flavor quark is through the value of  $L_{2D}$ .

## 4.6 Data, Monte Carlo, and Background

### 4.6.1 Data

The data used in this analysis is composed of three primary datasets, referred to as “0d”, “0h”, and “0i”, combined into  $955 \text{ pb}^{-1}$ . These consist of events with CEM, CMUP and CMX triggers, with separate storage for the electron (CEM) and muon (CMUP or CMX) triggered events. The divisions into “0d”, “0h” and “0i” reflect different periods in the operation of the CDF detector; information on the separate datasets, including the number of  $t\bar{t}$  candidate events passing all selection cuts<sup>4</sup>, is shown in Table 4.1.

### 4.6.2 Monte Carlo

In order to both construct and test analyses it is necessary to have large quantities of simulated signal data in which the true value of the underlying

---

<sup>4</sup>excluding an additional cut we designed which is dependent on our analysis calculation, see Sec. 6.2

Dataset	Date Range	$\text{pb}^{-1}$	# Evt. Selected
0d	Feb. '02 - Aug. '04	333	66
0h	Dec. '04 - Sep. '05	362	62
0i	Sep. '05 - Feb. '06	260	51

Table 4.1: *The three datasets, corresponding to different periods of Tevatron running, used in our measurement.*

quantity or quantities of interest (in our case, the top pole mass) is known, and simulated background data in which the background process is known. These events are referred to as “Monte Carlo” (MC) events, and they are created using sophisticated software packages which model the physics processes of interest. The signal events used to test the analysis were generated at tree level and then hadronized using Herwig 6.508 [31], while the background events used were constructed using Alpgen v1 [32] at the tree level but then hadronized with Herwig 6.508. On the other hand, many of the MC signal events used to study systematic effects in the measurement (see Chap. 9) were generated and hadronized using Pythia v6.2 [33], including the systematic related to the choice of Monte Carlo event used to construct the analysis.

Some added features of interest in the MC should be noted. With our Herwig signal events, in order to properly handle the decays of heavy flavor hadrons (those containing  $b$  and  $c$  quarks) the package QQ v9.1 [34] was used in conjunction with the hadronization and decay performed by Herwig. GEANT3 [35] was the software package used to describe the materials and geometry of the CDF detector in order to provide a framework for detector response to the jets of particles created; additional subsidiary packages were used to model the physics of specific subsystems. Also, the actual instantaneous luminosity and beam position during different runs (subdivisions of Tevatron stores defined for data acquisition) were used as parameters in the MC. The MC used in this analysis



was modeled using these parameters from only the 0d dataset period. This is a slight limitation in that, as muons in the miniskirt and keystone region of the CMX weren't triggered on for a portion of the initial runs comprising the 0d dataset, the signal MC we use doesn't have these muons either, even though datasets 0h and 0i, as well as a portion of 0d, do have them. In order to remain consistent, then, we removed the few data events which contained these types of muons from our sample. In order to ensure that the MC events accurately modeled the data events used in our analysis, we compared distributions of various event observables in for both MC and data; the results can be found in Appendix A and demonstrate that the MC is acceptable.

### 4.6.3 Background

Despite the selection cuts involved, a portion of candidate  $t\bar{t}$  events is expected to consist of events which are background (i.e., non- $t\bar{t}$ ). Background events (with the negligible exception of single top events) contain no information on the mass of the top quark, and their effect needs to be minimized in the measurement of the top mass. Therefore, in order both to optimize and to understand the performance of the analysis it is necessary to properly estimate the expected fraction of these events in the data sample as well as properly model them in MC. The expected number of different types of background events in the sample can be seen from Table 4.2; how these values were obtained will be described shortly. As can be seen from the table, the largest contributors to background are events in which a  $W$  decays leptonically in association with a set of partons which proceed to hadronize into high-energy jets (“ $W + \text{light}$ ” and “ $W + \text{heavy}$ ”) and events in which there is no  $W$ , but at least five jets, one of which is misidentified as a high- $p_T$  lepton (“non- $W$ ”). Other types of background – single top events as well as diboson processes such as the production of  $WW$ ,  $WZ$ , etc. – are predicted to be in the sample but in low quantities.

Event Type	# 1 Tag Evts.	# > 1 Tag Evts.
<b>non-W QCD</b>	<b><math>5.5 \pm 1.1</math></b>	<b><math>0.13 \pm 0.07</math></b>
W+light (mistag)	$9.5 \pm 1.6$	$0.65 \pm 0.32$
diboson (WW, WZ, ZZ)	$1.4 \pm 0.3$	$0.07 \pm 0.02$
<b>Sum of above 2</b>	<b><math>10.8 \pm 1.9</math></b>	<b><math>0.72 \pm 0.34</math></b>
W + $b\bar{b}$	$4.3 \pm 1.6$	$0.90 \pm 0.25$
W + $c\bar{c}$ , W + $c$	$2.9 \pm 1.0$	$0.13 \pm 0.07$
Single top	$0.6 \pm 0.1$	$0.0 \pm 0.0$
<b>Sum of above 3</b>	<b><math>7.8 \pm 2.7</math></b>	<b><math>1.03 \pm 0.32</math></b>
<b>Total Background</b>	<b><math>24.1 \pm 3.4</math></b>	<b><math>1.88 \pm 0.48</math></b>
<b>Total Events in Data</b>	<b>132</b>	<b>47</b>

Table 4.2: *Expected number of background events given the selection cuts described in this chapter. Events with more than 1  $b$  tag have a much lower proportion of background than events with only 1  $b$  tag.*

For this reason we model our background using only  $W$  + light,  $W$  + heavy and non- $W$  MC events, assuming the fraction of  $W$  + light events to include diboson events and the fraction of  $W$  + heavy to include single top events, as shown in the table. How background is dealt with in this analysis is described in Sec. 6.3.

The primary backgrounds are described as follows. “ $W$  + heavy” refers to those  $W$  + jets events in which actual production of heavy flavor quark(s),  $c$  or  $b$ , occurs; Monte Carlo with  $Wc\bar{c}$ ,  $Wb\bar{b}$  and  $Wc$  at the generator level (i.e., Feynman diagram level) is used to model these events. This type of background has a relatively low cross section, but a high efficiency. Conversely, “ $W$  + light” refers to those events in which only light quarks ( $u$ ,  $d$  or  $s$ ) are produced, but one of the jets gets mistagged as a heavy flavor jet. To calculate the rate at

which this occurs, a “mistag matrix” is used, which takes into account the probability a light jet would be tagged as a function of jet  $\eta$ ,  $E_T$ , and other parameters. Monte Carlo events with a  $W$  and light partons are used to model this type of background. Finally, in the case of “non- $W$ ” events, there is a much lower efficiency on this type of event than in the  $W + \text{jets}$  events given the selection cuts, but the cross section for multijet events without a  $W$  is very high. For this reason, non- $W$  events represent a significant contribution to the total background. Given the very low efficiency for non- $W$  events it is computationally impractical to model them with Monte Carlo; therefore non- $W$  events used for construction and evaluation of the analysis aren’t modeled, but rather are taken from data.

Like the other CDF top mass measurements of its generation and before, this analysis uses the expected number of background events calculated in top cross section measurements, as the selection cuts used are almost exactly the same<sup>5</sup>. Specifically, we use the background numbers from a CDF measurement of the top cross section using  $695 \text{ pb}^{-1}$  [46], but correct for the fact that this analysis did not require exactly four tight jets. A correction, then, is obtained from the most recent analysis to calculate the backgrounds using exactly four tight jets, a top cross section measurement performed with  $318 \text{ pb}^{-1}$  of data [47]; scale factors of 0.895 for events with 1  $b$ -tag and 0.736 for events with  $>1$   $b$ -tags are thus applied to the expected number of background events for our analysis. Additionally, due to the fact that we eliminate miniskirt and keystone muons for the reasons described earlier this chapter, further scale factors of 0.980 for 1  $b$ -tag and 0.968 for  $> 1$   $b$ -tags were calculated and applied as this removal is not done in [46]. Finally, the raw numbers of expected background events in the data sample are scaled up to reflect the increase in luminosity; a systematic

---

<sup>5</sup>As of July 2007, this is being replaced by a technique which allows any analysis to input its selection cuts and have its expected background numbers calculated

of 10% to the background fraction is applied in the measurement due to this scaling (see Chap. 9 for more on this).



## Chapter 5

### Signal Likelihood

## 5.1 The Full Formula

The analysis presented in this thesis falls in the category of matrix element analyses, which have become extremely popular in recent years. The foundation of the analysis is the same foundation that underlies all matrix element methods: a calculation of an event-by-event likelihood in the top pole mass,  $M_t$ . The way this is achieved is to integrate over a set of possible parton-level kinematics for the process of interest, i.e., the production of  $t\bar{t}$  pairs and their subsequent decay,  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$ . For every given parton-level kinematics, a weight (the integrand) is calculated; this weight is proportional to the square of the Standard Model matrix element for the process, as well as the probability given by the transfer functions that the measured jet momenta would have been produced by the partons ( $b\bar{b}q\bar{q}'$ ) with the assumed momenta. A simplified expression for the likelihood, then, is:

$$L(\vec{y} \mid M_t) = \int W(\vec{y} \mid M_t, \vec{x}) d\vec{x} \quad (5.1)$$

where  $\vec{x}$  is a vector of parton-level kinematic quantities,  $\vec{y}$  is the vector of measured quantities in the  $t\bar{t}$  candidate event,  $M_t$  is the top pole mass, and  $W(\vec{y} \mid M_t, \vec{x})$  is the aforementioned weight.

The full expression for the likelihood is more complicated, and is written as:

$$L = \frac{1}{N(M_t)} \frac{1}{A(M_t, \text{JES})} \sum_{perm=1}^{24} w_{perm} \int \frac{f(q_{z1})f(q_{z2})}{FF} \text{TF}(\vec{y} \cdot \text{JES} \mid \vec{x}) |M_{eff}(M_t, \vec{x})|^2 d\Phi(\vec{x}) \quad (5.2)$$

with  $L = L(\vec{y} \mid M_t, \text{JES})$ . A complex formula, this chapter will endeavor to motivate and explain its various pieces. Note that the likelihood is a function not just of  $M_t$  but also of “JES”, a scale factor which multiplies the jet momenta and is designed to compensate for the lack of knowledge of the calorimeter’s jet energy response. A very simple addition (and one not without precedent, see [10]), this will be shown in Chap. 7 to effectively convert the jet energy

systematic into a statistical error. Separate likelihoods are calculated for the  $4! = 24$  possible jet-parton permutations and summed; in general, the correct permutation will have a higher likelihood value than the others, and in this sense the method provides a natural permutation weighting. The other terms in the likelihood are as follows:

- $w_{perm}$  represents a weight assigned to the given permutation; it is described in Sec. 5.9
- $f(q_{z1})$  and  $f(q_{z2})$  are the CTEQ5L distribution functions for the incoming parton momenta [37]. Separate ones are used for gluons and quarks.
- $FF$  is the flux factor,  $1/\sqrt{|q_1||q_2|}$ , for the (assumed massless) incoming partons
- $|M_{eff}(M_t, \vec{x})|^2$  is the square of Standard Model matrix element with a modification; the matrix element is described in Sec. 5.4 and the modification is described in Sec. 5.5.
- $TF$  refers to the transfer functions, probability distribution functions for the jet momentum,  $|\vec{p}_j|$ , given the  $p_T$  of a parton; the transfer functions are described in Sec. 5.6.
- $d\Phi(\vec{x})$  is the differential of the particles' phase space; a discussion of the phase space of the integration can be found in Sec. 5.3.
- $N(M_t)$  is a term which normalizes the likelihood (i.e., it aims to establish  $\int L(\vec{y}) d\vec{y} = 1$ ) without regard for the selection cuts applied to  $\vec{y}$ ; this is described in Sec. 5.7.
- $A(M_t, \text{JES})$  is needed to correct for the selection cuts in order to have a properly normalized likelihood; this is described in Sec. 5.8.



The likelihood is calculated over the range  $M_t = 110$  to  $267.5$  GeV/c<sup>2</sup> in steps of  $2.5$  GeV/c<sup>2</sup>, and JES =  $0.85$  to  $1.15$  in steps of  $0.006$ . An interpolation is performed between these points such that the final step size is  $0.2$  GeV/c<sup>2</sup> in  $M_t$  and in  $0.002$  in JES.

## 5.2 Comments on the Technique

Matrix element methods are a powerful way to measure the mass of the top quark. For one thing, they take advantage of the knowledge represented in the Standard Model when measuring an event. Additionally, they are capable of accounting for the non-gaussian error on the measurement of jet momenta. Also, in general the likelihoods for events which are background will be wider than for those which are signal, and in this sense matrix element methods provide a useful event-by-event weighting given that individual event likelihoods are multiplied to produce a final likelihood for the sample.

Still, matrix element methods also can suffer from limitations: to date, all matrix element analyses have consisted of integrations only on a subspace of possible parton-level kinematics. Whereas in this analysis,  $\vec{x}$  is a 7-dimensional vector of kinematic quantities that gets integrated over, in principle it could be 22-dimensional. The reason for this is that, as there are 6 decay products in the  $t\bar{t}$  decay, it would take 6 momentum four-vectors - or 24 kinematic values - to completely describe the decay. However, the masses of both leptons are known<sup>1</sup>, reducing the number of imperfectly known kinematic values to 22. In fact, it should be noted that in a truly complete integral over the kinematics of the  $t\bar{t}$  production and decay process, the kinematics of the incoming partons also need to be integrated on. However, in our analysis, the incoming partons are taken to be massless with no transverse momentum in the rest frame of the

---

<sup>1</sup>To a precision satisfactory for this analysis!

$t\bar{t}$  system. As a result, their energies and z-momenta are completely determined by the energy and z-momentum of the six decay products, eliminating the need for the integration in question.

As a computational matter, integrating over a 22-dimensional space on an event-by-event basis is difficult. It is possible to reduce the dimensionality of the space by introducing kinematic assumptions into the integration. All matrix element analyses make (at least) the following assumptions, as we do:

- The quarks' masses are known
- The quarks' angles are known
- The lepton momentum is perfectly measured
- The mass of the neutrino and muon or electron is perfectly known

The hadronic  $b$  is taken to have a mass of  $4.95 \text{ GeV}/c^2$ ; the other quarks are taken to be massless, including the leptonic  $b$ . The leptonic  $b$  is massless as this only requires a fourth-order polynomial to be solved in the calculation of the kinematics rather than the eighth-order polynomial necessary if a nonzero mass is assumed. Naturally, some of these assumptions, such as the mass of the neutrino, are better than others, such as the mass of the quark. For this reason, the matrix element used in this analysis is adjusted,  $M(M_t, \vec{x}) \rightarrow M_{eff}(M_t, \vec{x})$ , to account for this. It should also be pointed out that while the integration outlined in this section is the foundation of the analysis, it is not the final word. A modification to the likelihood performed in order to account for the presence of background events in the sample will be described in Chap. 6, and the manner in which a measurement is extracted from these event-by-event likelihoods will be described in Sec. 7.1. The alteration of the  $t\bar{t}$  production and decay matrix element, as well as the background handling method, make the analysis presented in this thesis unique among matrix element analyses.

### 5.3 Phase Space Integration

The integration in Eq. 5.2 is performed over  $d\Phi(\vec{x})$ , rather than  $d\vec{x}$  itself. This is equivalent, of course, to a simple integration over the space of the kinematic variables,  $\vec{x}$ , but with an added factor of  $d\Phi/d\vec{x}$  in the integrand.  $d\Phi/d\vec{x}$  is the “phase space factor”, needed to account for the fact that more final quantum states are available to higher-energy particles due to the Lorentz contraction of volume at relativistic velocities [13]. Before expressing  $d\Phi$  as a function of the kinematic variables integrated on, however, first an expression for the differential phase space of the 6-body  $t\bar{t}$  decay needs to be written down using the four momenta of its particles:

$$d\Phi_6(q_1 + q_2; p_1, p_2, p_3, p_4, p_l, p_\nu) = \delta^4(q_1 + q_2 - \sum_{i=1}^6 p_i) \prod_{i=1}^6 \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i} \quad (5.3)$$

Here, the  $q_i$ ’s represent the four-momenta of the two incoming partons, and the  $p_i$ ’s represent the four-momenta of the four outgoing quarks in the decay, the  $p_\nu$  represents the four-momenta of the neutrino and  $p_l$  that of the lepton. Note that in the summation ( $\sum$ ) and product ( $\prod$ ) terms, indices 5 and 6 are used to label the  $\nu$  and  $l$ . Eq. 5.3 can be rewritten in terms of the quark’s angles and absolute momenta:

$$d\Phi_6 = \delta^4(q_1 + q_2 - \sum_{i=1}^6 p_i) \prod_{i=1}^4 \left( \frac{\rho_i^2 d\rho_i}{(2\pi)^3 2E_i} d\vec{\Omega} \right) \frac{d^3\vec{p}_e}{(2\pi)^3 2E_e} \frac{d^3\vec{p}_\nu}{(2\pi)^3 2E_\nu} \quad (5.4)$$

Here, the  $\rho_i$ ’s are the *magnitudes* of the quark momenta, and the  $d\vec{\Omega}$ ’s are the solid angles to which they point. In this form, the differential phase space is expressed as a function of kinematic variables, some of which, as described in the previous section, have their values explicitly assumed - the angles of the quarks and the momentum of the lepton. In an integral over this differential phase space, then, the angles of the quarks and the lepton momentum can be handled through delta functions, and these values will not affect the likelihood’s

dependence on the top mass or the JES. In practice, then, the portion of the differential phase space of interest is:

$$d\Phi_6 = \delta^4(q_1 + q_2 - \sum_{i=1}^6 p_i) \prod_{i=1}^4 \frac{\rho_i^2 d\rho_i}{(2\pi)^3 2E_i} \frac{d^3\vec{p}_\nu}{(2\pi)^3 2E_\nu} \quad (5.5)$$

This leaves seven variables to integrate on; however these are not the variables used in the integration, represented by  $\vec{x}$ . To properly calculate the phase space using the integration variables, we'll need to calculate the Jacobian between the integration variables and the variables in 5.5; in addition we'll need to rid ourselves of the delta function. To eliminate the delta function, we will want to integrate on the difference between the four-vector momentum of the outgoing particles and that of the incoming particles, choosing our variables as follows:

$$\begin{aligned} P_E &= E_{q_1} + E_{q_2} - \sum_{i=1}^6 E_i \\ P_x &= -\sum_{i=1}^6 p_{ix} \\ P_y &= -\sum_{i=1}^6 p_{iy} \\ P_z &= q_{1z} + q_{2z} - \sum_{i=1}^6 p_{iz} \\ \beta &= \log(\rho_1/\rho_2) \\ m_{W_{had}}^2 &= (p_1 + p_2)^2 \\ m_{t_{had}}^2 &= (p_1 + p_2 + p_3)^2 \\ m_{W_{lep}}^2 &= (p_4 + p_e + p_\nu)^2 \\ m_{t_{lep}}^2 &= (p_l + p_\nu)^2 \end{aligned} \quad (5.6)$$

where here, the transverse momenta and the masses of the incoming particles are neglected. In this case, we have it that  $q = (q_z, 0, 0, q_z)$ , or  $E_q = q_z$ , for the two incoming partons, and 5.5 becomes a function of nine variables: the magnitudes of the quark momenta, the neutrino momentum, and the z-momenta of the incoming partons. Of the integration variables shown in 5.6,

$m_{t_{had}}^2$  and  $m_{t_{lep}}^2$  are the squares of the masses of the top quarks on the hadronic and leptonic side of the decay,  $m_{W_{had}}^2$  and  $m_{W_{lep}}^2$  are the squares of the masses of the  $W$ 's on the hadronic and leptonic side of the decay, and  $\beta = \log(\rho_1/\rho_2)$  is the log of the ratio of the magnitudes of the light quark momenta from the hadronic side decay. The squared  $W$  and top masses are useful to integrate on as they account for most of the variation of the value of  $|M(M_t, \vec{x})|$  in the space of  $\vec{x}$ ; see Sec. 5.10 for further discussion.

We now need to calculate the Jacobian:

$$J = \frac{\partial(P_E, P_z, \beta, m_{W_{had}}^2, m_{t_{had}}^2, m_{W_{lep}}^2, m_{t_{lep}}^2, P_x, P_y)}{\partial(q_{1z}, q_{2z}, \rho_1, \rho_2, \rho_3, \rho_4, p_{\nu x}, p_{\nu y}, p_{\nu z})}$$

$$= \begin{pmatrix} 1 & -1 & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} \\ 1 & 1 & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} \\ 0 & 0 & a_{33} & a_{34} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{53} & a_{54} & a_{55} & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} \\ 0 & 0 & a_{73} & a_{74} & a_{75} & 0 & a_{77} & a_{78} & a_{79} \\ 0 & 0 & a_{83} & a_{84} & a_{85} & a_{86} & -1 & 0 & 0 \\ 0 & 0 & a_{93} & a_{94} & a_{95} & a_{96} & 0 & -1 & 0 \end{pmatrix}$$

Not all of the  $a_{ij}$ 's here need to be calculated; in fact, the Jacobian factorizes into three separate Jacobians for the incoming partons as well as the hadronic and leptonic side of the decay:

**Incoming partons :**  $|J_{in}| = \left| \frac{\partial(P_E, P_z)}{\partial(q_{1z}, q_{2z})} \right| = 2$

**Hadronic side of the decay :**  $|J_{had}| = \left| \frac{\partial(\beta, m_{W_{had}}^2, m_{t_{had}}^2)}{\partial(\rho_1, \rho_2, \rho_3)} \right| = (a_{33}a_{44} - a_{34}a_{43})a_{55}$

**Leptonic side of the decay :**  $|J_{lep}| = \left| \frac{\partial(m_{t_{lep}}^2, m_{W_{lep}}^2, P_x, P_y)}{\partial(\rho_4, p_{\nu x}, p_{\nu y}, p_{\nu z})} \right| = a_{66}a_{79} + (a_{67}a_{79} - a_{69}a_{77})a_{86} + (a_{68}a_{79} - a_{69}a_{78})a_{96}$

where we have

$$\begin{aligned}
a_{33} &= \frac{\partial \beta}{\partial \rho_1} = 1/\rho_1 \\
a_{34} &= \frac{\partial \beta}{\partial \rho_2} = -1/\rho_2 \\
a_{43} &= \frac{\partial m_{W_{had}}^2}{\partial \rho_1} = 2 (E_{W_{had}} \rho_1 / E_1 - \vec{p}_1 \cdot \vec{p}_{W_{had}} / \rho_1) \\
a_{44} &= \frac{\partial m_{W_{had}}^2}{\partial \rho_2} = 2 (E_{W_{had}} \rho_2 / E_2 - \vec{p}_2 \cdot \vec{p}_{W_{had}} / \rho_2) \\
a_{55} &= \frac{\partial m_{t_{had}}^2}{\partial \rho_3} = 2 (E_{t_{had}} \rho_3 / E_3 - \vec{p}_3 \cdot \vec{p}_{t_{had}} / \rho_3) \\
a_{66} &= \frac{\partial m_{t_{lep}}^2}{\partial \rho_4} = 2 (E_{t_{lep}} \rho_4 / E_4 - \vec{p}_4 \cdot \vec{p}_{t_{lep}} / \rho_4) \\
(a_{67}, a_{68}, a_{69}) &= \frac{\partial m_{t_{lep}}^2}{\partial \vec{p}_\nu} = 2 (E_{t_{lep}} / E_\nu \vec{p}_\nu - \vec{p}_{t_{lep}}) \\
(a_{77}, a_{78}, a_{79}) &= \frac{\partial m_{W_{lep}}^2}{\partial \vec{p}_\nu} = 2 (E_e / E_\nu \vec{p}_\nu - \vec{p}_e) \\
a_{86} &= \frac{\partial P_x}{\partial \rho_4} = -\sin \theta_4 \cos \phi_4 \\
a_{96} &= \frac{\partial P_y}{\partial \rho_4} = -\sin \theta_4 \sin \phi_4
\end{aligned} \tag{5.7}$$

The explicit formula for the Jacobian simply involves plugging in the  $a_{ij}$ 's shown in 5.7; in the interests of space the full, final expression is omitted. Our final phase space factor (ignoring constant terms which won't affect the likelihood) is:

$$\frac{d\Phi(\vec{x})}{d\vec{x}} = \frac{1}{|J_{in}| |J_{had}| |J_{lep}|} \frac{1}{E_\nu} \prod_{i=1}^4 \frac{1}{E_i} \tag{5.8}$$

## 5.4 The Matrix Elements

The matrix elements for the processes of concern, quark annihilation,  $q\bar{q}' \rightarrow t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$ , and gluon fusion,  $gg \rightarrow t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$ , are implemented for this analysis using the formalism of Kleiss and Stirling [38]. Strengths of the Kleiss-Stirling matrix elements include the fact that they take correlations between  $t\bar{t}$  production and decay into account, they allow the top and the  $W$  to have nonzero widths, and they account for the effects of a nonzero  $b$  mass; a matrix element calculation we previously worked with [39, 40] was not able to account for widths or  $b$  mass. The expected fraction of the production cross section of quark annihilation events vs. the cross section of gluon-gluon fusion events

is taken into account in our calculation; we take them to be their calculated values of 0.85 and 0.15 in the Standard Model at Next-to-Leading-Order (NLO), respectively [19]. It should be noted that, strictly speaking, the Kleiss-Stirling matrix elements model the  $t\bar{t}$  decay in the dilepton channel. However, as pointed out in [38], the Kleiss-Stirling matrix elements can be modified to accomodate the lepton + jets or all-hadronic decays through adding simple factors to account for the flavor and color of the decay products. As these factors are constants independent of the top pole mass and the kinematics of the decay, they can be discarded in the calculation of the likelihood.

There are four Kleiss-Stirling matrix elements, corresponding to the leading order quark annihilation process and the three gluon fusion processes which can occur up to NLO, shown in Fig. 2.1. These matrix elements are more fully discussed in Appendix B; there they are represented as products of Dirac spinors, and their presentation closely follows that of Kleiss and Stirling's paper. The final result is obtained by summation and squaring of the matrix elements, followed by summing and averaging over the spins of the incoming and outgoing particles.

## 5.5 $M(M_t, \vec{x}) \rightarrow M_{eff}(M_t, \vec{x})$ : The Effective Propagators

The integrand in Eq. 5.2 can be divided into two parts; one, the transfer functions, is supposed to provide the probability that the given parton kinematics,  $\vec{x}$ , would have produced the measured quantities in the detector,  $\vec{y}$ . The other contains terms whose product is supposed to be proportional to the probability that  $\vec{x}$  would have been produced in a  $p\bar{p}$  collision in the first place. The most significant determinants of this latter value are the masses of the top and  $W$  quarks in the decay. Using the matrix elements described in the

previous section, their effect on the probability is in the form of Breit-Wigner distributions for the  $m_W^2$  and  $m_t^2$  contained within the  $|M(M_t, \vec{x})|^2$  term. The Breit-Wigner distributions are simply the squares of the  $W$  and top propagators in the  $|M(M_t, \vec{x})|^2$ . If  $B(m_t^2)$  is the Breit-Wigner distribution for the squared mass of top, we have

$$B(m_t^2) \propto \frac{1}{(m_t^2 - M_t^2) + \Gamma M_t} \quad (5.9)$$

where  $M_t$  is the top pole mass and  $\Gamma$  is the top width, and similarly for the  $W$ .

However, there is an inconsistency: the kinematics which are solved for in the integration space of  $\vec{x}$ , described in Sec. 5.2, contain assumptions which are not entirely accurate. Given this fact, there's no reason to believe that the  $m_W^2$  and  $m_t^2$  which are reconstructed using the kinematic assumptions should possess Breit-Wigner distributions. And, in fact, they do not. However,  $|M(M_t, \vec{x})|^2$  is supposed to contain information on the probability of the production of a  $t\bar{t}$  decay with a particular  $\vec{x}$ . In order to make it more accurate, then, it makes sense to replace the Breit-Wigner distributions of  $m_W^2$  and  $m_t^2$  with distributions which reflect the reality of our assumed kinematics. We do just this in our analysis, and refer to the resulting matrix element with the term “ $M_{eff}(M_t, \vec{x})$ ”, for “effective” matrix element. How we calculate these new distributions for the effective matrix element, known as “effective propagators”<sup>2</sup>, will now be described.

### 5.5.1 Construction of Effective Events

The building block of our new distributions is a generator-level Herwig MC  $t\bar{t} \rightarrow l + \text{jets}$  event which we modify so that its kinematics obey our assumptions. Given this event, we remodel the top and  $W$  decays such that, while the

---

<sup>2</sup>This is really a misnomer, as the distribution of the particle mass squared is really the square of its propagator. Unfortunately, this nomenclature has become entrenched in the analysis!



momentum four vector of the decay system is preserved, the decay products adhere to our mass assumptions. With this modification, we take the four new partons and rotate them into the same directions as their counterparts in the original MC event. Then, we randomize the angle of the parton in such a way as to mimic jet angle resolution effects in the detector. The functions which form the basis of the randomization are parameterized by the  $p_T$  of the parton to account for the fact that higher-momenta partons have better measurements of their angles; the  $p_T$ -dependent angular resolution from MC events is shown in Fig. 5.1. Finally, the momentum of the charged lepton is randomized according to a gaussian centered at the initial value of the momentum, with a width determined by the expected detector resolutions:  $0.1 p_T^2$  % (muon case) or  $13.5/\sqrt{p_T}$  % (electron case).

The final-state partons and charged lepton of this modified generator-level Herwig MC event adhere to the very same kinematic assumptions made in the integration. To reconstruct the corresponding  $m_W^2$  and  $m_t^2$  values on the hadronic side, then, the invariant masses of the relevant partons simply need to be calculated. On the leptonic side, in order to calculate  $m_W^2$  and  $m_t^2$  we take the transverse momentum of that side of the decay,  $\vec{P}_{Tlep}$ , to be the difference between the original Herwig  $\vec{P}_{Ttt}$  and the sum of the transverse momenta of the three hadronic side partons just constructed,  $\vec{P}_{Thad}$ . There is still an ambiguity in the solution, however, as there is more than one neutrino's  $p_z$  solution. We take the value of the neutrino  $p_z$  to be that which provides  $m_t^2$  and  $m_W^2$  solutions closest to the original Herwig values, “closeness” here being defined by a  $\chi^2$ :

$$\chi^2 = \frac{(m_t^2 - m_{t_{orig}}^2)^2}{\sigma_t^2} + \frac{(m_W^2 - m_{W_{orig}}^2)^2}{\sigma_W^2} \quad (5.10)$$

where the  $\sigma$ 's are the products of the width and pole mass of the particles, and  $m_{t_{orig}}^2$  and  $m_{W_{orig}}^2$  denote the squared masses of the top and the  $W$  in the original Herwig event.

With the steps described in this section, a mapping has been created be-

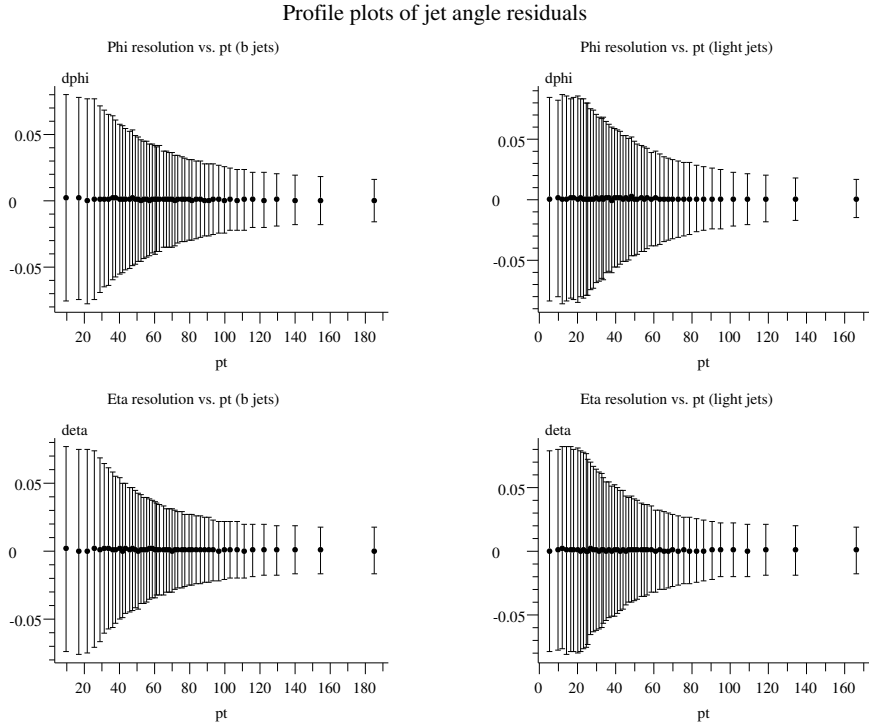


Figure 5.1: *Profile plots of differences between measured jet and true parton angles for  $b$  jets (left) and light jets (right),  $\phi$  (top) and  $\eta$  (bottom), as a function of jet  $p_T$ .*

tween the kinematics of a standard Herwig event and of an “effective” event whose kinematics should, on average, mimic the kinematics we solve for in our integration space given our assumptions.

### 5.5.2 Constructing the Effective Propagators

The effective events described in the previous section are used to construct the effective propagators; how this is done is very simple in the case of the leptonic-side propagator, somewhat less so for the hadronic-side propagator. The leptonic-side propagator is calculated simply by taking all available effective MC events, ranging in top mass from  $M_t = 150$  to  $M_t = 200$ , and forming a distribution of the events’  $m_W^2$  and the difference of the reconstructed and pole top masses,  $m_t - M_t$ . This latter approach is necessary as the  $m_t^2$  distributions have different means at different values of  $M_t$ ; it is a trivial matter to convert the calculated distribution to one in  $m_W^2$  and  $m_t^2$  during the integration when the top pole mass is assumed.

The hadronic-side effective propagator, on the other hand, is built with the following in mind: uncertainties on assumed kinematic quantities such as the quark masses and angles will have a greater or lesser effect on the uncertainties on the reconstructed top and  $W$  masses in an event depending on that event’s measured kinematics. This is more significant for the hadronic side of the decay rather than the leptonic side, due to there being three partons rather than just one. For this reason, unlike the effective propagators for the leptonic-side top and  $W$ , the effective propagators on the hadronic side are a function of the event’s measured kinematics. For a given event, the variances of  $m_W^2$  and  $m_t^2$ ,  $\sigma_{m_W^2}^2$  and  $\sigma_{m_t^2}^2$ , as well as their correlation,  $\rho_{m_W^2, m_t^2}$ , can be calculated. This is done using standard, first-order multivariate error propagation formulas given the  $\sigma_\eta$  and  $\sigma_\phi$  for each jet, as well as the error on the jet mass squared,  $\sigma_{m^2}$ , taken to be a constant  $241.7 \text{ GeV}^2/c^4$  for  $b$  jets and  $201.8 \text{ GeV}^2/c^4$  for light jets

based on Herwig MC jet mass distributions. It is possible to categorize events based on their calculated  $\sigma_{m_W^2}^2$ ,  $\sigma_{m_t^2}^2$  and  $\rho_{m_W^2, m_t^2}$ , and consequently construct separate distributions of  $m_W^2$  and  $m_t^2$  from effective events belonging to separate categories. When calculating the likelihood for an event, then, the event's  $(\sigma_{m_W^2}^2, \sigma_{m_t^2}^2, \rho_{m_W^2, m_t^2})$  is calculated and then the corresponding distributions are used as that event's hadronic-side effective propagators.

Plots of hadronic and leptonic side propagators are shown in Fig. 5.2.

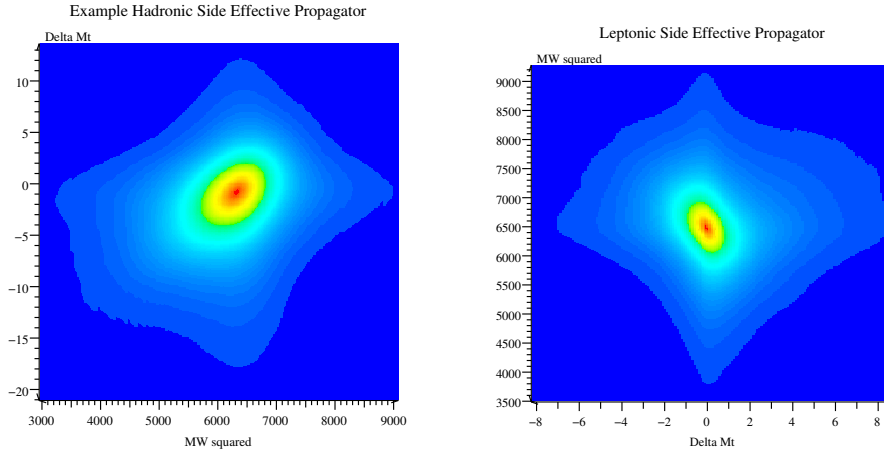


Figure 5.2: *The hadronic-side and leptonic-side propagators in  $m_W^2$  and  $m_t - M_t$ . In the integration,  $m_t - M_t$  is converted to  $m_t^2$  for an assumed  $M_t$ .*

## 5.6 Transfer Functions

### 5.6.1 Overview

If, for an assumed kinematics  $\vec{x}$ , the matrix element is meant to provide information on the probability that  $\vec{x}$  would have occurred in the first place, the transfer function is meant to provide the probability that given an  $\vec{x}$ , the quantities  $\vec{y}$  would have been measured in the detector. Specifically, our transfer functions are designed to calculate the probability that a parton with a given

$p_T$  would have produced the momentum of its resulting jet, corrected up to Level 5<sup>3</sup>; note that this is consistent with the jet momenta in our actual data events, which are also corrected to this level. Different transfer functions exist for the case of a  $b$  vs. a light quark, to account for differences in the two types of resulting jets. Additionally, separate transfer functions have been constructed for four different regions of the detector, to account for the effect of cracks between calorimeters and differing amounts and types of material which jets with different  $\eta$  values shower through. These regions consist of  $0 < |\eta| < 0.15$ , corresponding to a region where the two halves of the detector meet,  $0.15 < |\eta| < 0.85$ , the fiducial region of the central calorimeters,  $0.85 < |\eta| < 1.4$ , a region containing the cracks between the central, wall and plug calorimeters, and  $1.4 < |\eta| < 2.0$ , consisting entirely of the plug calorimetry.

### 5.6.2 Construction

The method by which the transfer functions are constructed involves creating distributions from MC events of the ratio of a given jet's absolute momentum to the energy of its underlying parton,  $\rho_j/E_p$ , binned by the  $p_T$  of the underlying parton. A function<sup>4</sup> parameterized by  $p_T$  is then fit simultaneously to all the distributions. In Fig. 5.3 distributions for light quarks can be seen in the  $0.15 < |\eta| < 0.85$  region; note from the plot that the function is fit only to jets which would have passed the selection cut, but is extrapolated and normalized over all values of  $\rho_j/E_p$ . The jets used in these distributions are taken from the Herwig signal MC with the standard selection cuts, along with an additional requirement that the event actually involve a  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  decay. All four jets in the event are used in transfer function fits; the decision as to which parton matches which jet is made by calculating  $\chi^2$  values for all possible jet-parton

---

<sup>3</sup>See Sec. 4.4 for a description of the jet momenta correction levels

<sup>4</sup>Johnson's functions, described in Appendix C.1

matches and choosing the permutation with the best fit. The  $\chi^2$  is defined simply as

$$\chi^2 = \sum_{i=1}^4 \left( \frac{\eta_i^j - \eta_i^p}{\sigma_\eta} \right)^2 + \left( \frac{\phi_i^j - \phi_i^p}{\sigma_\phi} \right)^2 \quad (5.11)$$

where the resolutions on  $\eta$  and  $\phi$ ,  $\sigma_\eta$  and  $\sigma_\phi$ , are both taken to be 0.05.

### 5.6.3 Transfer Function Study

As a sanity check of the transfer functions, a “toy” likelihood calculation of the top and  $W$  masses on the hadronic side of the  $t\bar{t}$  decay is performed. This calculation uses only the transfer functions, as well as a pair of priors for the top and  $W$  momenta,  $G(\vec{p}_t)$  and  $F(\vec{p}_W)$ . In this manner, an attempt has been made to isolate the performance of the transfer functions from the rest of the standard machinery of the measurement technique. Only the best match between the four underlying partons and the four jets is used. The measurement is taken to be the mass at which the product of the individual event likelihoods (the “joint likelihood”) assumes its maximum value; this is the same technique as is used in the full analysis measurement (as described in Chap. 7), though without the bias and error calibrations. A full explanation follows.

#### $W$ Mass Measurement

The formula for the event-by-event likelihood in the  $W$  mass cross-check is as follows:

$$L(p_{j1}, p_{j2} \mid M_W) \propto \int \frac{p_1^2}{E_1} \frac{p_2^2}{E_2} \frac{E_W}{p_W^2} TF(p_{j1} \mid p_{T1}) TF(p_{j2} \mid p_{T2}) f(p_W, \cos \theta_W) \frac{d\beta}{|J_W|} \quad (5.12)$$

Note that in this formula, constant terms have been dropped; these terms, of course, will not matter for the calculation. For notational simplicity, the  $p$ ’s here are meant to denote the absolute values of the vector 3-momenta of the particles. Additionally,

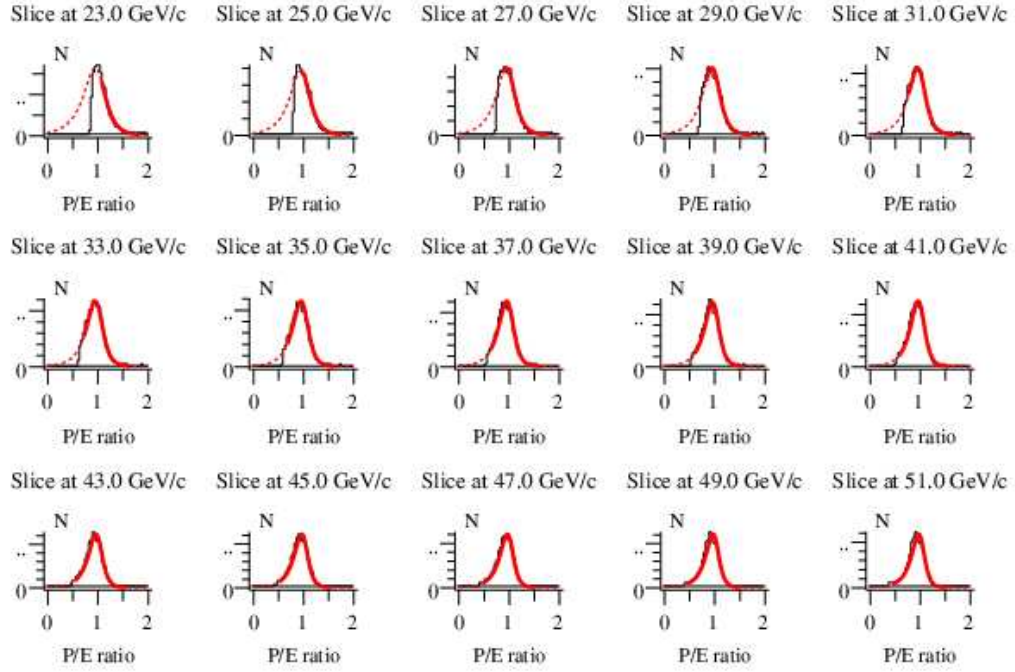


Figure 5.3: *Fitted transfer functions for light quarks in the fiducial region of the central calorimeters ( $0.15 < |\eta| < 0.85$ ). The solid line shows the region of the distribution to which the function was fitted, while the dotted line represents the region of extrapolation.*

- Subscript  $W$  corresponds to the  $W$  boson; subscripts 1 and 2 correspond to the two light partons from the  $W$  decay, and subscripts j1 and j2 correspond to their L5-corrected jets
- The integration variable, “ $\beta$ ”, has the same definition as in the main likelihood integral (Eq. 5.2).
- $TF$  is the transfer function
- $f(p_W, \cos \theta_W)$  is the prior for the  $W$  momentum; note that it is not a function of the  $\phi$  of the  $W$  momentum
- $|J_W|$  is the Jacobian to account for the transformation of variables needed given the integration over  $\beta$

The derivation of this formula is given in Appendix D, including a discussion of the construction of the prior. In the calculation,  $\beta$  is integrated over the range  $[-1.95 + \log(p_{j1}/p_{j2}), 1.95 + \log(p_{j1}/p_{j2})]$  in steps of 0.1. The likelihood is calculated over a  $W$  mass from 20 to 160 GeV/c<sup>2</sup> in steps of 0.5 GeV/c<sup>2</sup>, and is then convoluted with a Breit-Wigner with a width equal to that of the  $W$ , 2.1 GeV/c<sup>2</sup>, to account for the fact that width was neglected in Eq. 5.12.

## Top Mass Measurement

The full formula for the integration performed to calculate the top mass, analogous to Eq. 5.12 for the  $W$  case, is:

$$L(p_{j1}, p_{j2}, p_{j3} \mid M_t) \propto \int \frac{p_1^2}{E_1} \frac{p_2^2}{E_2} \frac{p_3^2}{E_3} T F_1 T F_2 T F_3 \frac{E_W}{p_W^2} \frac{E_t}{p_t^2} f(p_W, \cos \theta_W) g(p_t, \cos \theta_t) \frac{d\beta}{|J_t|} \quad (5.13)$$

where here,  $TF_i \equiv TF(p_{ji} \mid p_{Ti})$ . The assumptions and simplifications underlying the form of this integral proceed in a completely analogous manner to that of the  $W$  mass measurement, and are also covered in Appendix D, along with the construction of the top momentum prior,  $g(p_t, \cos \theta_t)$ . For this integration, the



$\beta$  range and step size is the same as that for the  $W$  mass integration described in the previous section. The width of the  $W$  is again ignored, and the value of the  $W$  mass for the event is taken to be its value measured using Eq. 5.12. The top mass is calculated over a range of 80 to 280 GeV/c<sup>2</sup> in steps of 0.5 GeV/c<sup>2</sup>, and similarly to the case of the  $W$  mass integration, a convolution is performed with a Breit-Wigner in which  $\Gamma = 1.5$  GeV/c<sup>2</sup> to account for the fact the the top width is neglected in the integration.

## Result

The techniques described above were applied to a sample of Herwig  $t\bar{t}$  events with a true top mass of 178 GeV/c<sup>2</sup> and a  $W$  mass of 80.4 GeV/c<sup>2</sup>. Calculating the likelihoods of all the events in the sample which pass the standard set of cuts and multiplying them, we obtain the joint likelihoods for the  $W$  and the top mass shown in Fig. 5.4. The measured mass of the  $W$  is found to be 80.4 GeV/c<sup>2</sup>, and that of the top to be 180.8 GeV/c<sup>2</sup>. It appears that, while the reconstructed top mass has a somewhat positive bias, it is fair to say that the transfer functions pass this “sanity check”.

## 5.7 Normalization

In order for  $L(\vec{y} \mid M_t, \text{JES})$  to properly be considered a likelihood, it needs to be normalized in its space of measured quantities,  $\vec{y}$ :

$$\int L(\vec{y} \mid M_t, \text{JES}) dy = 1 \quad (5.14)$$

In a sense, the title of this section is a misnomer, as it will cover only the component of our likelihood,  $1/N(M_t)$  in Eq. 5.2, which normalizes the likelihood *without* regard to the selection cuts. For a full normalization, however, the selection cuts need to be taken into account; this is covered in Sec. 5.8.

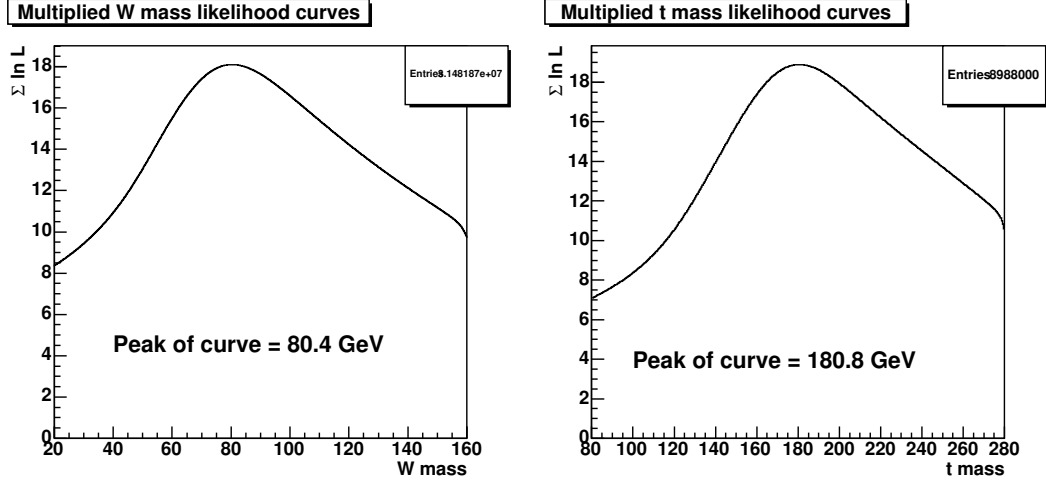


Figure 5.4: *Transfer function cross check results for hadronic  $W$  and top mass using fully simulated Herwig MC.*

### 5.7.1 Normalization Without Effective Propagator Modification

The normalization is calculated in two steps: the first step calculates the normalization ignoring the effects of our effective propagators, while the second step includes this effect. In the first step, an approximate integral over the parton-level kinematic space is performed. Note that here, the kinematic assumptions of our integration are *not* made, specifically meaning the partons do not have to obey our mass and angle assumptions; this space shall be denoted with  $\vec{x}'$ , as opposed to  $\vec{x}$  which is the space of kinematics which *do* obey our assumptions. The integral is performed as follows: points are generated in the space of  $t\bar{t}$  kinematics using Breit-Wigner distributions for  $W$ s and tops (assuming a given top pole mass) and the phase space density for the decay particles. At each point, a weight is calculated with the  $f(q_z)$ 's,  $FF$  and  $\Phi(\vec{x}')$  described in Sec. 5.2, along with  $|M^*(M_t, \vec{x}')|^2$ , denoting the  $t\bar{t}$  production and decay matrix element *excluding* the propagators. Summing the weights calculated at every point provides the integral; in Fig. 5.5 it can be seen that there

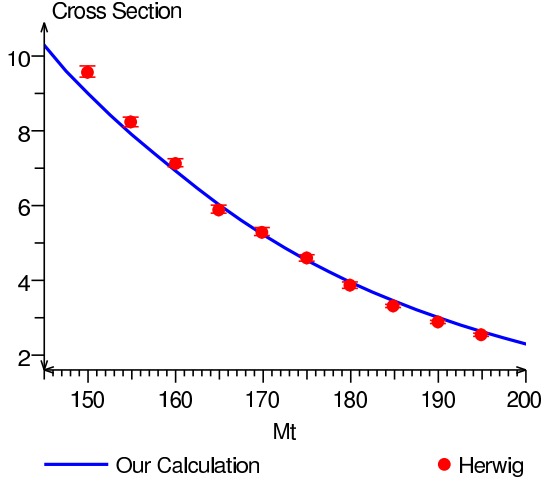


Figure 5.5: *Our normalization, before the modification for the effective propagators, compared to the Herwig cross section after an appropriate scaling.*

is reasonable agreement between the cross section as calculated in Herwig and the cross section calculated through our integration. A scale factor has been applied in the comparison as constant factors necessary for a true cross section calculation were omitted in our approach; these constant factors do not affect the likelihood as they are not top mass dependent. The difference between our result and the Herwig result can be attributed to the absence of radiation in our model.

### 5.7.2 Addition of the Effective Propagator Normalization

In order to modify  $N(M_t)$  to account for the effective propagators, we in essence modify the integral for a given top pole mass described in the previous section by calculating the average ratio of the value of the integrand when using the kinematic assumptions to the value of the integrand without those assumptions. This hinges on the approximation that there is a one-to-one map-

ping between a point in the phase space of kinematics without assumptions,  $\vec{x}'$ , to a point in the phase space of kinematics with assumptions,  $\vec{x}$ . In fact, the mapping is many-to-many given that we randomize the angles of the partons and the lepton momentum as described in Sec. 5.5.1; however to first order this can be used, and as our correction turns out to be small in the  $M_t$  region of interest, we stop there.

Given the one-to-one approximation and the fact that the  $m_W^2$  and  $m_t^2$  distributions on the hadronic and leptonic sides of the decay define the density of points in the kinematic phase space in this calculation, we have:

$$\Gamma_{eff}(\vec{x})d\vec{x} = \Gamma(\vec{x}')d\vec{x}' \quad (5.15)$$

where  $\Gamma_{eff}$  is a function which is the product of the densities of the four effective propagators, and  $\Gamma$  is the product of the densities of the four Breit-Wigners.

We can write the quantity we wish to calculate, the normalization given our kinematic assumptions, as:

$$N(M_t) = \int R(\vec{x})\Gamma_{eff}(\vec{x})|M^*(M_t, \vec{x})|^2 d\vec{x} \quad (5.16)$$

Here,  $|M^*(M_t, \vec{x})|^2$  is the matrix element squared without the product of the four distributions (i.e.  $\Gamma_{eff}(\vec{x}) \cdot |M^*(M_t, \vec{x})|^2 = |M_{eff}(M_t, \vec{x})|^2$ ), and  $R(\vec{x})$  is the rest of the integrand (composed of the incoming parton distribution functions and phase space factors). Eq. 5.16 can be rewritten as

$$\int \frac{R(\vec{x})\Gamma_{eff}(\vec{x})|M^*(M_t, \vec{x})|^2 d\vec{x}}{R(\vec{x}')\Gamma(\vec{x}')|M^*(M_t, \vec{x}')|^2 d\vec{x}'} R(\vec{x}')\Gamma(\vec{x}')|M^*(M_t, \vec{x}')|^2 d\vec{x}' \quad (5.17)$$

Given the equality in Eq. 5.15, Eq. 5.17 becomes:

$$\int \frac{R(\vec{x})|M^*(M_t, \vec{x})|^2}{R(\vec{x}')|M^*(M_t, \vec{x}')|^2} R(\vec{x}')\Gamma(\vec{x}')|M^*(M_t, \vec{x}')|^2 d\vec{x}' \quad (5.18)$$

which is simply the average value of

$$\frac{R(\vec{x})|M^*(M_t, \vec{x})|^2}{R(\vec{x}')|M^*(M_t, \vec{x}')|^2} \quad (5.19)$$

This quantity is calculated from generator-level Herwig events. For a given event, this is done by taking the initial kinematic quantities to be that event's  $\vec{x}'$ , modifying the event's kinematics to obey the standard assumptions to obtain that event's  $\vec{x}$  (as described in Sec. 5.5.1), and then calculating the value expressed in Eq. 5.19. By taking the average of this value for all events with a given  $M_t$  and doing this over the range of  $M_t$ 's for which we define the normalization, a correction factor which accounts for the effective propagators is calculated. This correction factor is shown in Fig. 5.6; as expected, it is uniformly less than zero as  $\vec{x}$  contains kinematic assumptions not contained in the Standard Model. When this correction factor multiplies the initial normalization we calculate without regard to the effective propagators, described at the start of this section, we obtain our final  $N(M_t)$ .

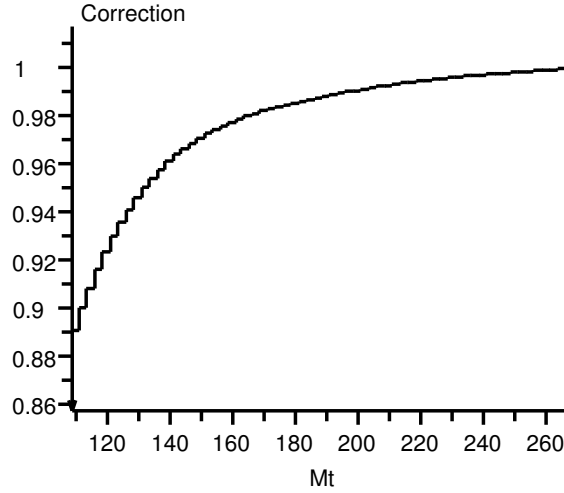


Figure 5.6: *The normalization correction applied to account for our use of effective propagators.*

## 5.8 Acceptance

The term in the likelihood shown in Eq. 5.2 which accounts for the effect of the selection cuts on the likelihood normalization is the acceptance,  $1/A(M_t, \text{JES})$ . All other factors disregard the selection cuts, as they are either parton-level quantities or the transfer functions, which, while fit only to jet momenta which pass selection, nonetheless are normalized without respect to the selection (Sec. 5.6). In order to account for this fact, we calculate the acceptance over the same range of  $M_t$  and JES over which the likelihood function is defined and employ it in the denominator of the likelihood calculation.

To calculate  $A(M_t, \text{JES})$ , we create ensembles of generator-level Herwig events which are adjusted to fit our kinematic assumptions (i.e., effective events), exactly as described in Sec. 5.5.1. However, we take an additional step in that the four partons in the event have their momenta randomized using the the very transfer functions we constructed; the resulting four-vectors can be considered “proto-jets”, given that the transfer function matches up parton  $p_T$  to jet momentum. In this way, events are created which are meant to mimic events which perfectly obey all the assumptions in our integration. Ensembles of these events are created at the same masses for which the likelihood is calculated,  $M_t = 110$  to  $267.5 \text{ GeV}/c^2$  in steps of  $2.5 \text{ GeV}/c^2$ . To mimic the effect of different true JES values we simply divide the proto-jet momenta by the assumed JES value. With ensembles of events, then, defined at every  $M_t$  and JES value at which the likelihood is defined, we calculate an acceptance by applying cuts meant to emulate the cuts applied to real events described in Chap. 4. These cuts are:

- The four proto-jets must have an  $|\eta| < 2$  and an energy  $> 21 \text{ GeV}$  (a cut of  $21 \text{ GeV}$  on an L5 corrected jet energy is approximately equivalent to a cut of  $15 \text{ GeV}$  at L4).
- The lepton must have a  $P_T > 20 \text{ GeV}$  and an  $|\eta| < 1$ .

- The proto-jets must be separated by a minimum  $\Delta R$  of 0.5 (this is because, in a real event, jets closer than this in  $\eta$ - $\phi$  space would be merged.)
- The lepton must be separated from each of the proto-jets by a minimum  $\Delta R$  of 0.5 (this should account for the lepton isolation cuts described in Secs. 4.2 and 4.3).
- $\cancel{E}_T$  must be at least 20 GeV. In a real event,  $\cancel{E}_T$  is the magnitude of the measured transverse energy of the event; the remaining energy is considered “missing” in that the  $p\bar{p}$  system is taken to have no transverse energy. The measured transverse energy is mimicked by summing transverse momenta of the four proto-jets and the lepton.

The resulting acceptance is shown in Fig. 5.7. An obvious question that arises is why we calculate our own acceptance rather than take the acceptance simply from applying our selection cuts to fully simulated MC events. The reason is that the purpose of the acceptance is to insure that the likelihood is normalized with respect to  $\vec{y}$ , the measured quantities in the detector. In our likelihood calculation, we make assumptions, among which are that we have a  $t\bar{t} \rightarrow l + \text{jets}$  event, that the physics process proceeds via the lowest order diagrams possible (e.g., without extra radiation), and that the jets selected originate with the four partons from the decay. However, these assumptions are clearly not satisfied by real events, and as a consequence calculating an acceptance using full simulation MC events ignores the details of our likelihood calculation.

## 5.9 Permutation weighting

For each of the 24 possible jet-parton permutations for which a likelihood is calculated, a weight  $w_{perm}$  is calculated which attempts to account for the

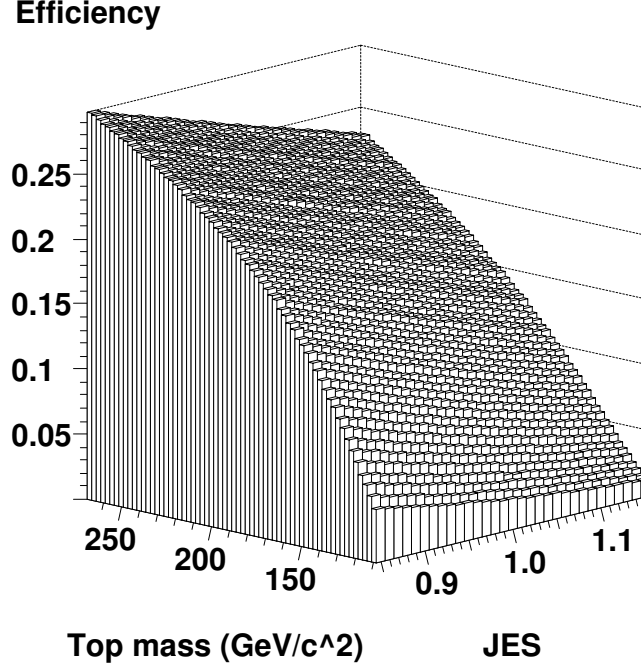


Figure 5.7: *The acceptance applied to the denominator in the likelihood function.*

probability that the observed jets would have been tagged given the assumed underlying partons. Using the tag efficiencies calculated by the CDF High  $p_T$   $b$ -Tag Group as a function of jet  $E_T$  and  $\eta$  for both  $b$  jets and mistagged light jets [41], we fit a 5<sup>th</sup> degree polynomial to these four sets of efficiencies in order to create smooth functions of the tag probabilities as a function of  $E_T$  and  $\eta$ . These fits are shown in Fig. 5.8 ; the probability,  $P_{tag}$ , that a light or  $b$  quark would have produced a tagged jet is simply the product of the relevant functions in  $\eta$  and  $E_T$ . Note that light quarks consist of  $u$ ,  $d$ ,  $s$  quarks and exclude  $c$  quarks; the  $P_{tag}$  for a  $c$  quark is simply taken to be  $0.22 \cdot P_{tag}$  for a  $b$  quark. For the  $i^{th}$  considered permutation, the weight for that permutation,



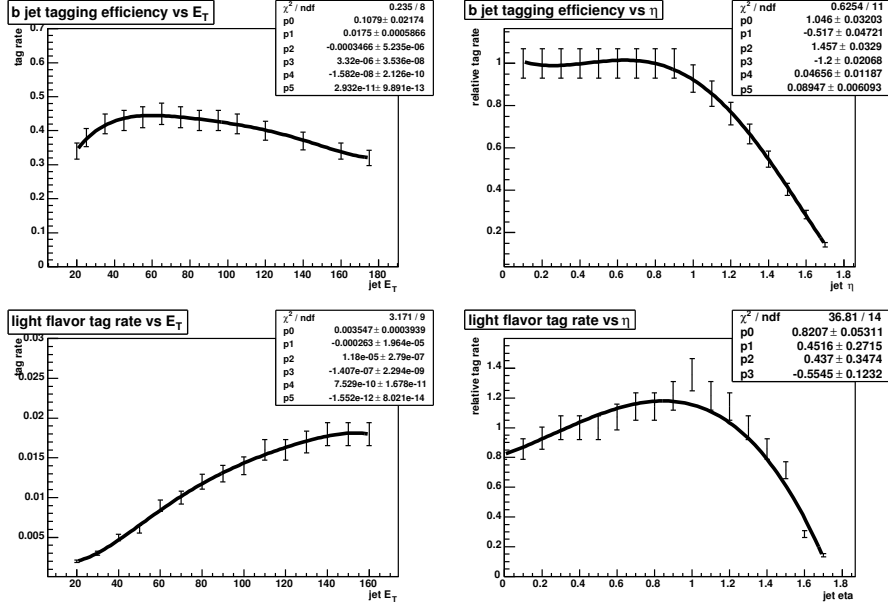


Figure 5.8: 5<sup>th</sup> degree polynomial fits to  $b$  tag efficiencies as a function of  $\eta$  and  $p_T$  for both  $b$  quarks and light quarks.

$w_i$ , is the product of four terms: the calculated  $P_{tag}$ 's for the tagged jets and  $(1 - P_{tag})$  for the untagged jets. It is assumed that the  $W$  decays half the time to a  $u$  and a  $\bar{d}$ , and half the time to a  $c$  and an  $\bar{s}$ ; this is significant due to the different manner in which the charm and light quark tag efficiencies are calculated.

## 5.10 Integration details

Integration over the kinematic space,  $\vec{x}$ , in the likelihood calculation is done numerically. A grid is set up with a fixed number of points to sample on the axes corresponding to  $m_W^2$  and  $m_t^2$  on the hadronic and leptonic sides of the decay, as well as for  $\beta = \log(\rho_1/\rho_2)$ ; additionally, a fixed number of points are used to sample the vector transverse momentum of the  $t\bar{t}$  system, as can be seen in Fig. 5.9. The numbers of sample points used for each variable can be found in

Table 5.1; they were determined through empirical studies designed to optimize the speed and accuracy of the integration. Taking the product of Table 5.1 entries, it can be seen that there are 338625 points in the integration space, each of which requires a kinematic solution for the integrand for each assumed top pole mass. However, this is computationally quite difficult, and as a result we calculate the integral using only a subset of these 338625 points. The choice of subset is designed to sample all regions of the space; this subset is determined by the “Sobol sequences” [42], capable of generating quasirandom numbers with coverage over the entire space. Only approximately one out of every 2000 points was actually used in the integration, though higher-frequency sampling was found to have little or no effect on the resolution of the measurement.

Not yet addressed has been the spacing of the points in the integration space. For the case of  $\beta$ , the spacing was simply equidistant in  $\beta$ . However, for the other variables, a choice was made to space the points equidistant in probability given their expected distributions. This is reflected in the choice of the  $t\bar{t}$  transverse momentum integration grid shown in Fig. 5.9, the width of the grid determined from values found in ensembles of Herwig MC signal events. Equidistant probability spacing is also used for the  $m_W^2$  and  $m_t^2$  variables, given their effective propagators. The reason for the selection of points equidistant in probability is that, given a fixed number of points with which to sample a space in an integral, the most accurate calculation can be performed with a spacing chosen in this manner.

variable	# pts	spacing (prob/flat)
$m_{W_{had}}$	25	prob
$m_{W_{lep}}$	5	prob
$m_{t_{had}}$	3	prob
$m_{t_{lep}}$	3	prob
$\beta$	17	flat
$\vec{P}_T$	301	prob

Table 5.1: *Here we have the # of points which sample the given variable's dimension, and info on whether the spacing of these points is equidistant in probability or constant in the variable*

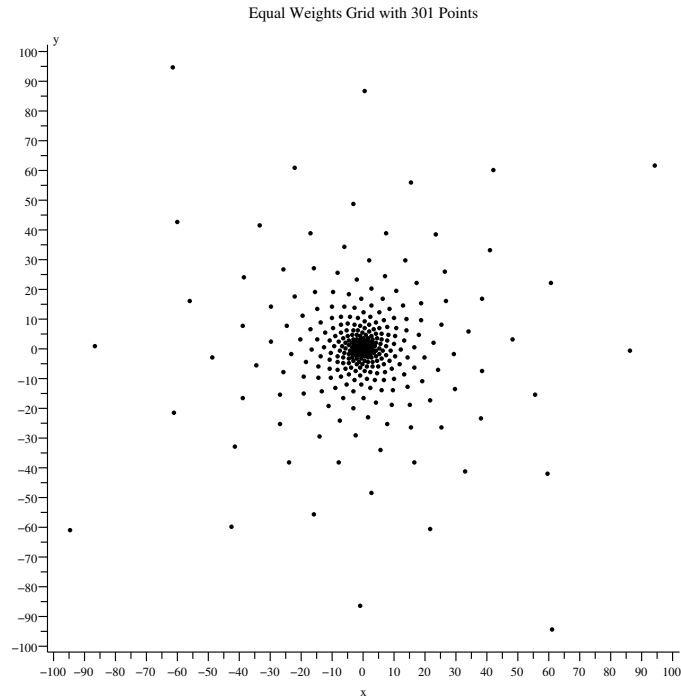


Figure 5.9: *The points of the  $t\bar{t}$  transverse momentum integration grid, chosen to reflect the values found in MC.*



## Chapter 6

### Undesirable Events

## 6.1 Introduction

A reality of this analysis is that, no matter how tight the selection cuts applied to candidate events for the  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  process, it is never possible to obtain a sample of signal-only events with 100% certainty. Based on the numbers of expected background events in the sample shown in Table 4.2, using standard selection cuts for  $t\bar{t} \rightarrow l + \text{jets}$ , approximately 15.5% of our 179 candidate events are expected to be background events. Given that these events carry no information on the top quark mass, they can only be expected to hurt our measurement<sup>1</sup>.

Interestingly, there is also a class of  $t\bar{t}$  signal events which, rather than improving the resolution of our measurement, actually hurt it. In our MC, these are  $t\bar{t}$  events which either do not undergo the expected decay of  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$ , or which have a poor best match between the final four partons of the decay and the four tight jets associated with them; collectively they are referred to as “bad” signal, with the remaining signal labeled “good” signal. The technical meaning of “poor best match” is based on the definition of a  $\chi^2$  for a set of jet-parton matches described in Eq. 5.11. If, for an event, the lowest  $\chi^2$  obtained by a jet-parton match permutation is greater than 150, or if the difference between the second-best and best  $\chi^2$  values is less than 200, the event is considered to have a poor match. The reasons for a poor match are typically that at least one of the tight jets doesn’t directly result from the fragmentation and hadronization of any of the four partons, but rather results from initial- or final-state radiation during the  $t\bar{t}$  decay process (e.g., from a gluon radiating off one of the incoming partons in the production of the  $t\bar{t}$ ). This type of event comprises approximately 30% of our signal MC, and its effects are shown in Sec. 7.2. If a  $t\bar{t}$  event fails to decay into the expected channel or if one or more of the tight jets in the event do not originate from one of the expected partons, the signal likelihood

---

<sup>1</sup>And they do; see Sec. 7.5 for studies on this.

calculation machinery described in the previous section is ill-equipped to extract information on the top quark mass. This is perhaps not too surprising in light of the fact that the calculation is based on the premise that (1) the physics process under consideration is  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  and (2) all four tight jets carry a good deal of information on the kinematics of their parent partons.

## 6.2 Signal Likelihood Cut

There is a very simple selection cut, in addition to the cuts described in Chap. 4, which effectively discriminates against both background and bad signal: a cut on the maximum value of an individual event's log likelihood. Specifically, we require that the maximum value of the log of an event's 2-d likelihood curve,  $\log L_{max}(\vec{y} \mid M_t, \text{JES})$ , be greater than 6. In Fig. 6.1, distributions of  $\log L_{max}$  for background, bad signal, and good signal are shown. The signal consists of Herwig events at  $M_t = 175 \text{ GeV}/c^2$ , and the background consists of  $W + \text{light}$  and  $W + \text{heavy}$  MC and non- $W$  data events, combined in the proportions expected based on the selection cuts described in Chap. 4. Clearly, background and bad signal have a tail in the region of low  $\log L_{max}$  values unlike good signal events. With the cut placed at 6 for both the 1- and >1-tag events, only  $\sim 5\%$  of good signal events were eliminated, while around a quarter of bad signal events and 40% of background was eliminated. Studies revealed that a cut at 6 for both the 1 tag and the >1-tag case provides the best resolution for the measurement. The efficiencies for the  $M_t = 175 \text{ GeV}/c^2$  case are shown in Table 6.1; these values change by only a couple of percent over the mass range  $M_t = 160 \text{ GeV}/c^2$  to  $M_t = 190 \text{ GeV}/c^2$ . With application of this cut to the data sample, the number of events is reduced from 179 to 149; additionally, the resulting numbers of expected background and signal events in the sample after the likelihood cut, calculated from the values in Table 4.2 and the efficiencies

of the various event types under the cut, are shown in Table 6.2.

Table 6.1: *Efficiency of the selection cut  $\log L_{max} > 6$  at  $M_t = 175 \text{ GeV}/c^2$*

Event Type	1-tag	>1-tag
Good signal	94.7%	94.1%
Bad signal	73.7%	80.2%
Background	63.1%	57.5%

Table 6.2: *Expected numbers of background and signal in data after likelihood cut.*

Event Type	1-tag	>1-tag
Signal	92.7	39.8
W + heavy flavor	5.0	0.6
W + light flavor	6.4	0.4
non-W	3.8	0.1

## 6.3 Background Handling

### 6.3.1 Modified Event Likelihood

Despite the application of the likelihood cut, it can be seen in Table 6.2 about 11% of the sample remains background. Given that the calculation of the likelihood explicitly assumes a  $t\bar{t}$  event, it is necessary to find a technique to modify the likelihood to also account for the existence of this background. One approach is to use matrix elements not just for signal but for the different types of background as well; however, given the difficulty in composing a matrix



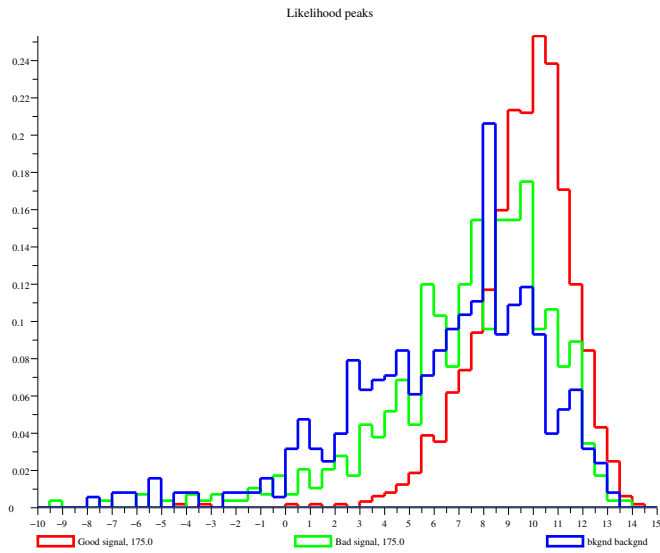
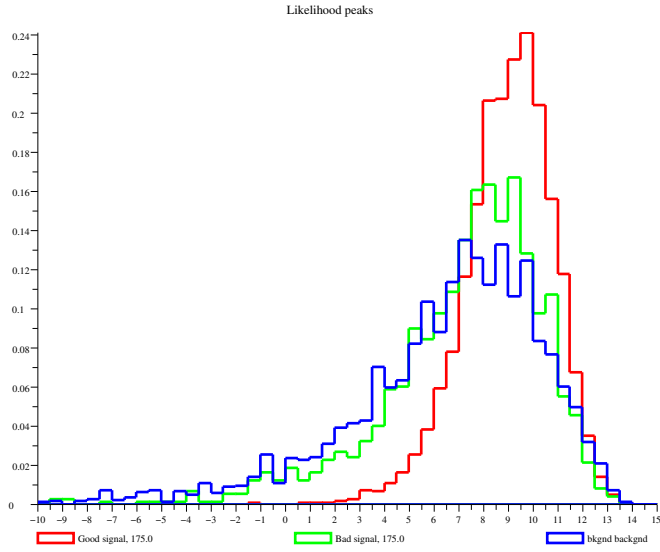


Figure 6.1: Distribution of maximum  $\log$  likelihood values for good signal (red), bad signal (green) and background (blue) in the 1-tag (top) and  $>1$ -tag (bottom) cases.

element to cover the various processes resulting in “non-W” events, we took an alternative approach.

We observe the following: in order to combine the likelihood curves calculated from an ensemble of  $N$  events for a measurement, a joint likelihood is formed, through multiplying the likelihood curves or, equivalently, through summing the logs of the likelihood curves:

$$\log L_{total}(M_t, \text{JES}) = \sum_{i=1}^{i=N} \log L_i(\vec{y}_i \mid M_t, \text{JES}) \quad (6.1)$$

Hypothetically, if we knew the average value of the background log likelihoods in our data sample,  $\overline{\log L_{bg}(M_t, \text{JES})}$ , as well as the number of background events in the sample,  $N_{bg}$ , we could completely eliminate the effect of background on the measurement. This average background log likelihood can be expressed as

$$\overline{\log L_{bg}(M_t, \text{JES})} = \frac{1}{N_{bg}} \sum_{j=1}^{j=N_{bg}} \log L_j(\vec{y}_j \mid M_t, \text{JES}) \quad (6.2)$$

where  $j$  runs over the likelihoods of the background events in the sample. Furthermore, if we were able to calculate a per-event probability that a given event were background,  $f_{bg}$ , and the probability were normalized such that

$$\sum_{i=1}^{i=N} f_{bg_i} = N_{bg} \quad (6.3)$$

then we could modify  $L_{total}$  to eliminate the effect of backgrounds in the sample

$$\log L_{final}(M_t, \text{JES}) = \sum_{i=1}^{i=N} (\log L_i - f_{bg_i} \overline{\log L_{bg}}) \quad (6.4)$$

Here,  $L_{final}$  is equivalent to the log likelihood we would obtain if we only combined the signal events in the sample.

It is with this in mind that we calculate  $\overline{\log L_{bg}}$  by summing the log likelihoods of our different background events in their expected proportions of Table 6.2. To account for background, then, for each event, we subtract off this

shape from the original likelihood curve weighted by  $f_{bg}$ . In the next section, the calculation of  $f_{bg}$  will be discussed;  $f_{bg}$  is calculated with an assumption of the number of background events in the sample taken from Table 6.2.

However, there is a problem: unlike the ideal example just described, clearly our calculated  $\overline{\log L_{bg}}$  will not be perfect, and  $N_{bg} \cdot \overline{\log L_{bg}}$  will *not* be the same log likelihood as the sum of the background log likelihoods in the data. Therefore, the effect of the background on the measurement will not be completely eliminated<sup>2</sup>. We attempt to compensate for this effect by reducing the fluctuations of a given sample's true  $\overline{\log L_{bg}}$  relative to the  $\overline{\log L_{bg}}$  we construct and use in the analysis. This is done by averaging each event's likelihood with the uniform distribution,  $U(M_t, \text{JES})$ , in effect “flattening” the likelihood. Flattening background events should reduce the error due to the aforementioned fluctuations; however, flattening signal events will increase the error as the resolution of a given signal event's likelihood will be reduced. As we don't know which events are background and which are signal, the extent to which we flatten an individual event is made a function of its calculated  $f_{bg}$ ; therefore, the flattening is expressed as:

$$\log L \Rightarrow \log[L(1 - f_{bg}) + f_{bg}U] \quad (6.5)$$

It is important to note that the log likelihoods used to construct the average background curve,  $\overline{\log L_{bg}}$ , also are subject to this flattening.

As a test to explore the tradeoff of errors obtained through flattening the likelihood curves, we experimented with a parameter,  $\kappa$ , which could modify the calculated background fraction used in the flattening,  $f_{bg} \rightarrow \kappa f_{bg}$ , and therefore increase or reduce the amount of flattening. In practice, we found no improvement in the error by using  $\kappa$ , so the final log likelihood of a sample

---

<sup>2</sup>A systematic error is calculated for this effect, see Chap. 9

with background handling included is:

$$\log L_{final}(M_t, \text{JES}) = \sum_{i=1}^{i=N} [\log\{L_i(M_t, \text{JES})(1 - f_{bg_i}) + f_{bg_i}U(M_t, \text{JES})\} - f_{bg} \overline{\log L_{bg}(M_t, \text{JES})(1 - f_{bg}) + f_{bg}U(M_t, \text{JES})}] \quad (6.6)$$

### 6.3.2 Calculation of a Background Probability, $f_{bg}$

#### Overview

In order to calculate a probability that a given event is a background event, we use both the initial estimate of the fraction of background taken from Table 6.2 as well as an event observable, hence referred to as “ $q$ ”, which has different distributions for signal and background events. “ $q$ ” will be defined in the next section; the focus here will be on the straightforward manner in which  $q$  is used to calculate  $f_{bg}$ .

First, we create histograms of  $q$  for signal MC at  $M_t = 175 \text{ GeV}/c^2$  as well as for  $W + \text{heavy}$  and  $W + \text{light}$  MC and non- $W$  data events. Events which have passed the primary selection requirements outlined in Chap. 4 as well as the likelihood cut described in earlier this chapter are used. These histograms are created separately for 1- and >1- tag events, and have their areas normalized to the expected fractions of background and signal in the data; these are shown in Fig. 6.2<sup>3</sup>. The separation of the 1- and >1-tag events is an attempt to do a better job of background handling given the very different expected background fractions between the two types of event.

For a given event, we calculate its value of  $q$ , take the corresponding bin of the signal and background histograms, and then simply calculate

$$f_{bg} = B(i_q)/(S(i_q) + B(i_q)) \quad (6.7)$$

where  $i_q$  is the  $i^{th}$  bin of the histogram which corresponds to the value of  $q$ ,

---

<sup>3</sup>similarly, we calculate  $\overline{\log L_{bg}}$  separately for the 1- and >1- tag cases

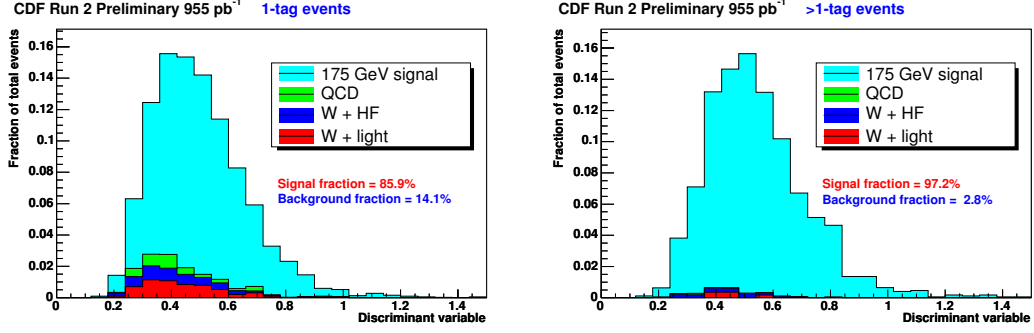


Figure 6.2: Distributions of the discrimination variable,  $q$ ; the histograms for signal and backgrounds are scaled to reflect the expected fractions of signal/backgrounds in the sample.

$B(i_q)$  is the sum of the background histograms in that bin, and  $S(i_q)$  is the value of the signal histogram in the bin. If the value of  $q$  falls outside the range over which the histograms are defined, then  $f_{bg} = 0$ .

## The Discrimination Variable

“ $q$ ” is defined with two properties in mind. First, and most importantly, it should be independent of the top pole mass and the JES, so that the same calculation of  $f_{bg}$  can apply over all values of  $M_t$  and JES used in the likelihood. Additionally, of course, it must also possess different distributions for signal and background.  $q$  is in fact a linear combination of event observables, so that the coefficients of the observables can be tuned to satisfy the first requirement. These event observables are defined as follows:

- The aplanarity  $A = \frac{3}{2}Q_1$ , where the  $Q_i$  are the eigenvalues  $Q_1 < Q_2 < Q_3$  of the momentum tensor  $\Theta_{ab} = \sum_i p_a^i p_b^i / \sum_i |\vec{p}_i|^2$ ; here  $a$  and  $b$  represent the coordinates  $x, y, z$  and  $i$  is the jet index.
- $D_R = \Delta R_{ij}^{\min} \cdot \min(p_z^{(i,j)}) / p_T(l^\pm)$ , i.e., the minimum separation  $\Delta R_{ij}^{\min}$  in  $\eta$ - $\phi$  space between a jet pair weighted with the ratio of the smaller  $p_z$  of

the jet pair and the charged lepton  $p_T$ .

- $H_{\text{TZ}} = \sum_{i=2}^4 |p_{\text{T}}^{(i)}| / (\sum_{i=1}^4 |p_z^{(i)}| + |p_z^{(l\pm)}| + |p_z^{(\nu)}|)$ , the ratio of the scalar sums of transverse (numerator) and longitudinal (denominator) momenta of decay objects in the event. The numerator contains all jets except the highest energy jet. The denominator sums all jets, the charged lepton and the neutrino. The smaller  $|p_z^{(\nu)}|$  solution given by the kinematic equation for the leptonic  $W$  decay (assuming  $M_W = 80.4 \text{ GeV}/c^2$ ) is taken.

It is noteworthy that all of these observables are measures not so much of energetic quantities in an event (as, for example, the sum of the jet  $E_T$ 's would be) but rather of the shape of the event. The reason for this is that, while energy-based variables can take advantage of the typically higher transverse energies of a signal event's decay products than those of a background event for discrimination, these variables are also, as one would expect, strongly correlated with  $M_t$ . On the other hand, due to the higher transverse energies the decay products in signal events, the geometry of these decay products in a signal event is more spherical than the products of a background event. Geometrically-based variables such as the ones we used, therefore, can discriminate between signal and background; however, as energy-based quantities are found both in the numerator and the denominator of these observables, they are much less dependent on  $M_t$  and JES than purely energy-based observables.

Distributions of all three observables can be found in Fig. 6.3, and it is apparent that each observable possesses different distributions for signal events compared to background events which pass the standard selection cuts. However,  $A$  and  $H_{\text{TZ}}$  have a clear dependence on the top mass, though only on the order of a few percent over a range of many 10's of GeV in the top pole mass, as can be seen in Fig. 6.4. Similarly, a dependence of all three variables on the jet energy scale can be seen in Fig. 6.5. It is noteworthy that the three observables do not have dependences on  $M_t$  and JES with the same slope; it

is this fact we take advantage of when we combine the observables in a linear combination. Studies indicated that the best way to minimize the dependence of our discrimination variable on  $M_t$  and JES was to alter the coefficients of the three component variables such that our final observable is:

$$q = \left( \frac{3}{4}A + \frac{1}{14}D_R + \frac{1}{2}H_{TZ} \right) \times \frac{1.4}{1.55} \quad (6.8)$$

where the  $\frac{1.4}{1.55}$  term is used so that  $q$  typically assumes values from 0 to 1.

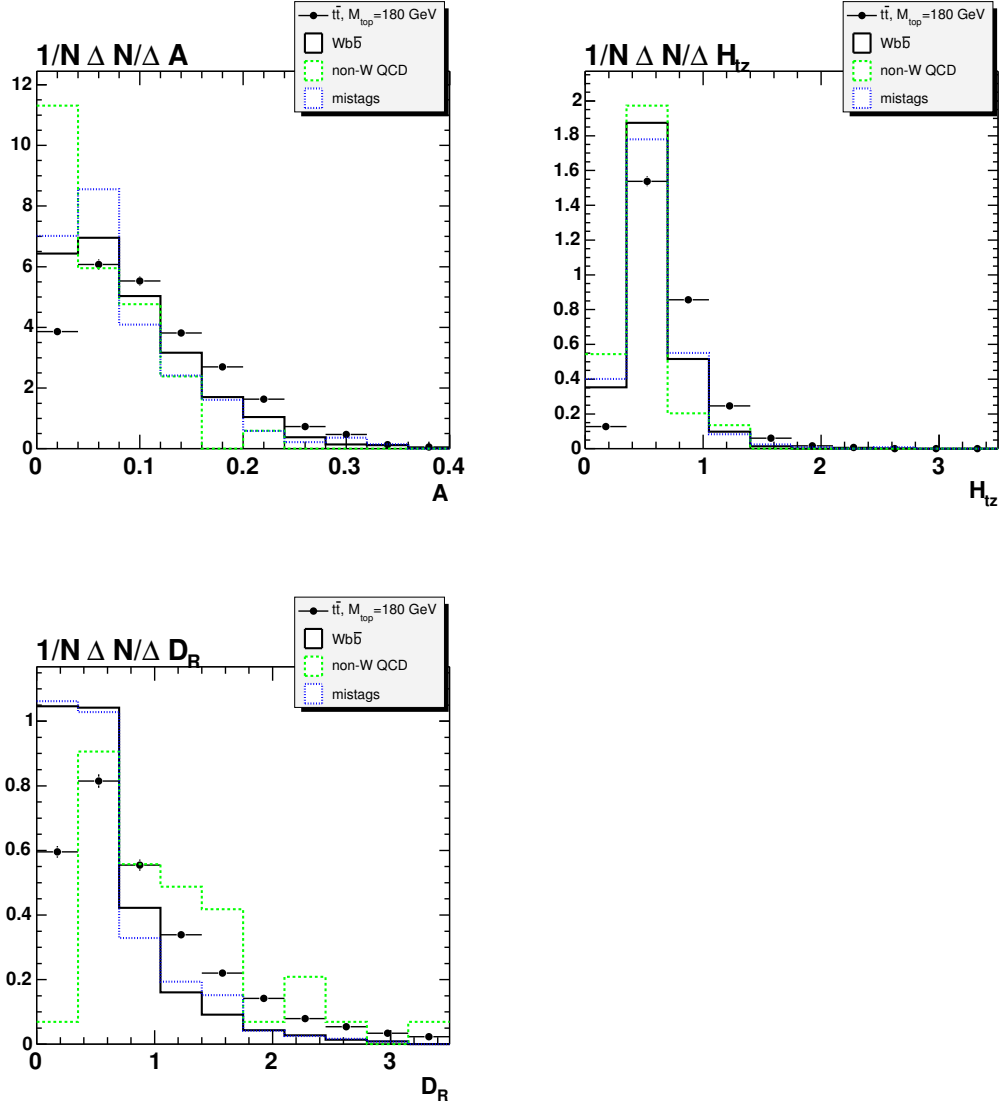


Figure 6.3: Distributions of aplanarity,  $H_{TZ}$  and  $D_R$ . Signal is at  $M_t = 180$   $\text{GeV}/c^2$ .



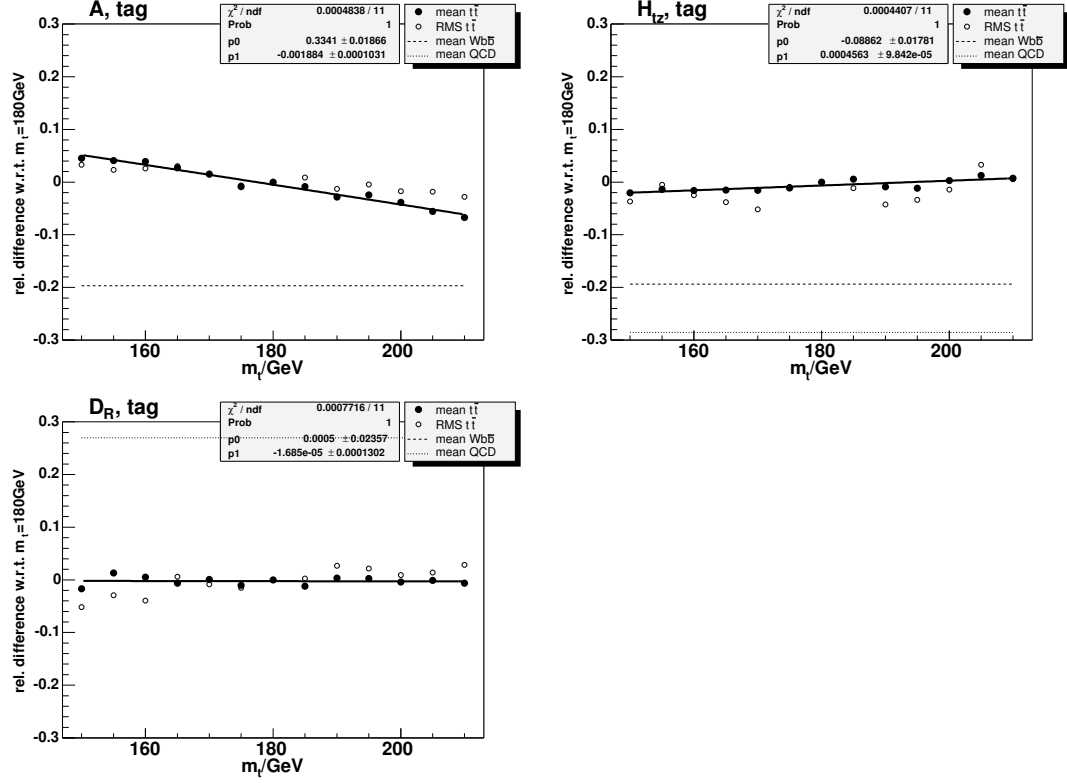


Figure 6.4: Bias of mean (black markers) and RMS (white markers) of distributions of  $A$ ,  $H_{TZ}$  and  $D_R$  for signal MC at different masses with respect to the mean and RMS of the distributions at  $M_t = 180 \text{ GeV}/c^2$ . The solid lines are fits to the biases on the mean, the dotted lines represent the bias on the mean of the two types of background under study.

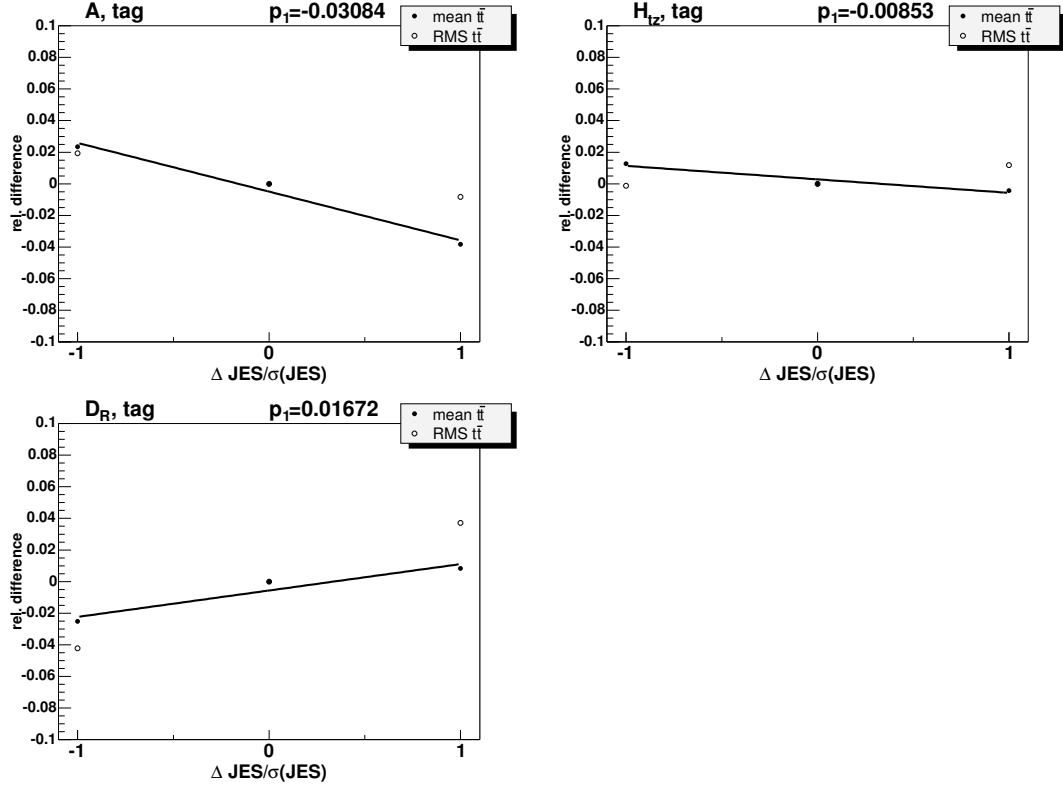


Figure 6.5: With the same format of the plots shown in Fig. 6.4, the plots here show the effect of shifting the jet energy corrections by  $+1$  and  $-1$   $\sigma$



## Chapter 7

### Tests of the Method

## 7.1 Performing a Measurement

When performing a measurement on a sample of events a sum of the logs of the individual event likelihoods is performed to obtain the final, joint log likelihood of the sample, as shown in Eq. 6.1. The manner in which we extract a measurement from this joint likelihood is quite straightforward. We take a profile of the 2-d log likelihood along the mass axis to obtain a 1-d log likelihood as a function of  $M_t$  only:

$$\log L_{total_{prof}}(M_t) = \max_{j \in \text{JES}} \log L_{total}(M_t, j) \quad (7.1)$$

With this 1-d log likelihood, then, the measured value of the mass,  $M_{meas}$ , is simply the location of the maximum log likelihood:

$$L_{total_{prof}}(M_{meas}) \geq L_{total_{prof}}(M_t) \quad (7.2)$$

and the quoted positive and negative errors on the measurement,  $\sigma_+$  and  $\sigma_-$ , are those mass shifts for which the log likelihood has a value 0.5 less than that of the maximum:

$$\log L_{total_{prof}}(M_{meas} - \sigma_-) = L_{total_{prof}}(M_{meas} + \sigma_+) = L_{total_{prof}}(M_{meas}) - 0.5 \quad (7.3)$$

– in other words, the log likelihood is simply treated as the log of a Gaussian, which should be a reasonable assumption given the number of events available in the data sample. In practice, we average the error to achieve a single quoted error,  $\sigma$ , on the measurement:

$$\sigma = \frac{1}{2}(\sigma_+ + \sigma_-) \quad (7.4)$$

## 7.2 Overview of Pseudoexperiments

Before performing the measurement on the data, it is necessary to ensure that the analysis will provide an unbiased measurement with an accurately

estimated error. In order to do this, we run the analysis on samples of MC signal and background events in order to determine any calibrations that will need to be applied to the initial measurement  $M_{meas}$ , and its quoted error,  $\sigma$ , in the data. These samples are drawn from ensembles of available MC events and are meant to mimic the sample of data events.

Running the analysis on a sample of MC events is known as a “pseudoexperiment”, or “PE”. For an ensemble of MC signal events with a known  $M_t$  and JES, 2000 PEs are run. To construct a sample for a PE, an average number of events of a given type is specified, and the actual number of events of that type in the sample is Poisson-fluctuated about that value. The types of event used in the PEs of this chapter are signal, or one of the three types of background described in Chap. 4 ( $W + \text{heavy}$ ,  $W + \text{light}$ , or non- $W$ ).

The main quantities of interest when running PEs on MC with a known  $M_t$  and JES are the bias, the error, and the pull width. These are defined as follows:

**Bias:** The difference between the true value and the mean of the 2000 PE measurements,  $\langle M_{meas} \rangle - M_t$

**Error:** The RMS of the 2000 PE measurements,  $\sqrt{\langle (M_{meas} - \langle M_{meas} \rangle)^2 \rangle}$

**Pull Width:** The RMS of each PE’s pull, where the pull is defined as the ratio of the bias and the individual PE’s error,  $(M_{meas} - M_t)/\sigma$ . If an analysis perfectly estimates its own errors, this should be unity.

These quantities have errors assigned to them, represented as error bars in the plots shown in this chapter. In the case of the bias, this error is simply taken to be the RMS of the distribution ( the “Error” described in the list) times the square root of the number of unique PEs which can be constructed given the number of signal events in the MC ensemble:

$$\sigma_{bias} = \sqrt{\langle (M_{meas} - \langle M_{meas} \rangle)^2 \rangle} \cdot \sqrt{\frac{N_{total}}{N_{PE}}} \quad (7.5)$$

where  $N_{total}$  is the number of signal MC events in the ensemble, and  $N_{PE}$  is the average number of signal events/PE.

Unlike in the case of the bias, it is possible to get a more accurate measure of the pull width and error through resampling the ensemble of MC events, i.e., running more than the number of unique PEs given the ensemble. In order to find the error on these quantities, then, we took an empirical approach. We took the ensemble of signal MC events at  $M_t = 175$ , dividing it into 8 equal sub-ensembles, and then running 2000 PEs with 1/8 the standard number of events/PE using signal MC taken exclusively from a single sub-ensemble. Doing this for each of the eight sub-ensembles, we obtained eight pull widths and eight errors; the RMS of these quantities came out to be 0.07 in the case of the pull width and 0.19 GeV/c<sup>2</sup> in the case of the error, which we assigned as the error on the pull width and the error. This was done only for the fully realistic case described in Sec. 7.4; however, these errors are also used for other, signal-only PEs, as the quantities themselves (pull width, etc.) are of greater interest than the errors on the quantities.

## 7.3 Signal-Only PEs

As a test of the signal integration described in Chap. 5, PEs were run using only signal MC events; the aspects of the analysis described in Chap. 6 designed to handle undesirable events - i.e., the background handling and the likelihood cut - were excluded from the analysis machinery. Specifically, two types of signal PEs were run: one involved using only the “good” signal events described in Chap. 6, and the other involved using all signal events, good and bad. In both cases, 179 events/PE were used; the good signal only PE results are shown in Fig. 7.1, and the good + bad signal PE results are shown in Fig. 7.2. It can be seen that in the case of the good signal results, the bias and the pull width

are both fairly reasonable - the bias is only a few tenths of a GeV, and the pull width is only a few percent above unity. However, when bad signal events aren't explicitly excluded from the PEs, performance degrades noticeably - the bias is now about 3 GeV/c<sup>2</sup>, the pull width is 30% above unity, and the value of the slope fitted to the errors at  $M_t = 172$  GeV/c<sup>2</sup> is about 50% higher than in the case of the good signal only PEs. The conclusion to be drawn from these results is that the likelihood formulated in Eq. 5.2 performs quite well when the  $t\bar{t}$  events it works with are in the assumed decay channel ( $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$ ) and the four tight jets in the event indeed come from the quark's hadronization and not unwanted radiation. However, it not only fails to extract information on the top mass from  $t\bar{t}$  events which don't adhere to these assumptions, its resolution is in fact *hurt* by them. Fortunately, these events are at least partially eliminated through the likelihood cut described in Sec. 6.2.

## 7.4 Fully Realistic PEs

In order to find the expected bias with which to shift the top mass measured from the data as to produce an unbiased measurement, as well as the expected pull width with which to multiply the initial error of the measurement to correctly estimate the error, it is necessary to run PEs which are as realistic as possible. In constructing these PEs, we required that the average number of events of the four major types be taken from Table 6.2; therefore the PEs consisted of an average of 148.8 events: 107.9 1-tag events and 40.9 >1-tag events. Additionally, unlike in the case of the signal-only PEs, these PEs were run using the full analysis machinery, i.e., with the likelihood cut and the background handling used to deal with undesirable events. The plots of the results are shown in Fig. 7.3. It can be seen that there is an expected bias in the measurement of 1.2 GeV/c<sup>2</sup> which is independent of the true top mass.



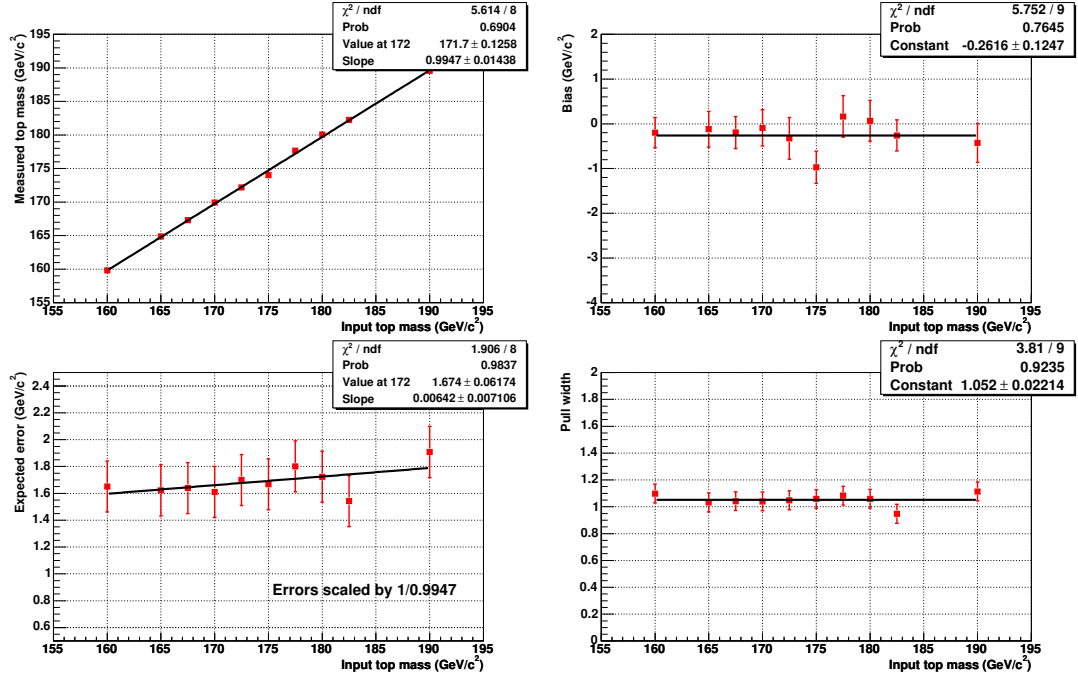


Figure 7.1: Results of PEs run with 179 good signal events each. Clockwise, from upper left:  $\langle M_{meas} \rangle$  vs.  $M_t$ ,  $\langle M_{meas} \rangle - M_t$  vs.  $M_t$ , pull width vs.  $M_t$ , and RMS of  $M_{meas}$  vs.  $M_t$ .

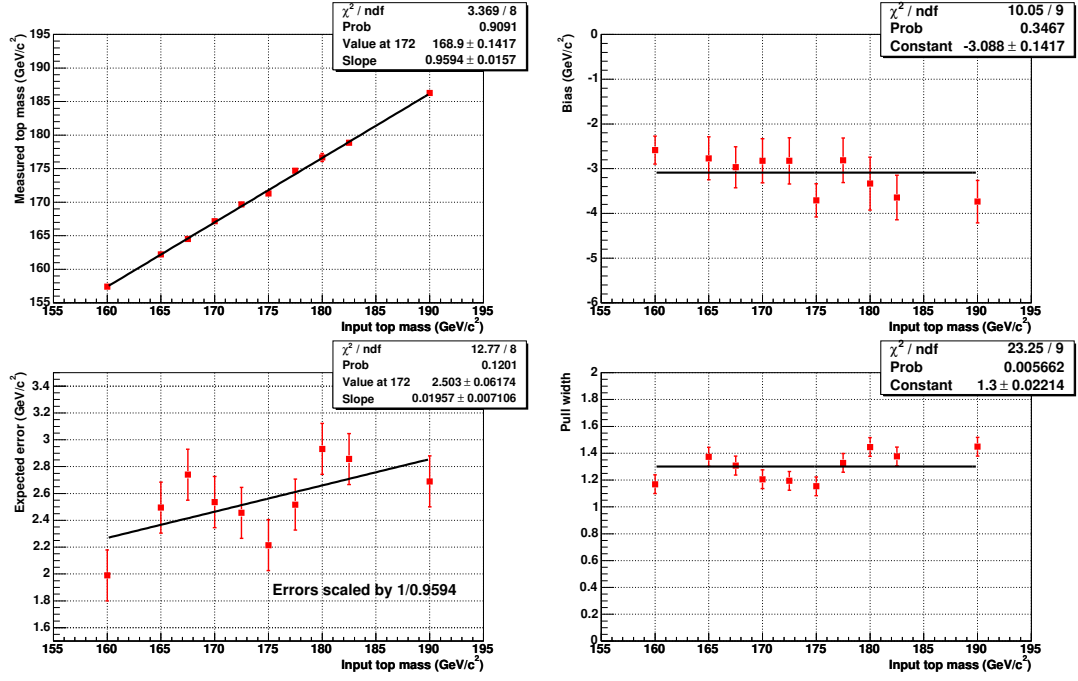


Figure 7.2: Results of PEs run with 179 signal events, good and bad. Plot format same as in Fig. 7.1

Additionally, the pull width is expected to be 1.22. These values are used to calibrate the actual data measurement described in the next chapter. Also of note is the expected error:  $2.5 \text{ GeV}/c^2$  at  $M_t = 172 \text{ GeV}/c^2$ .

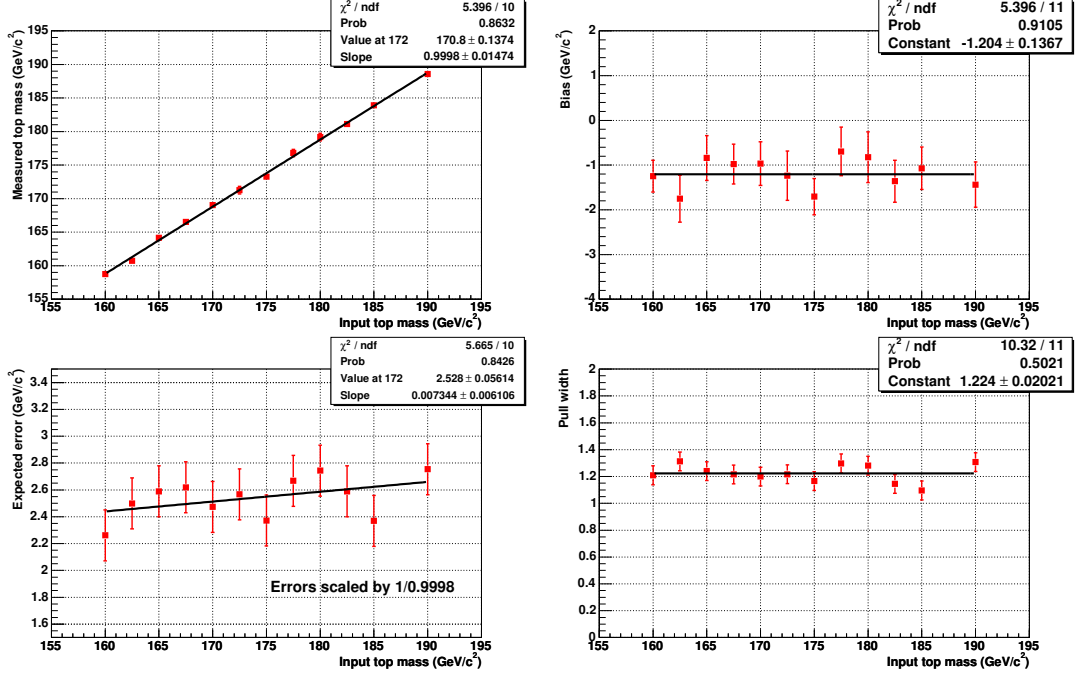


Figure 7.3: The results of fully realistic PEs. Plot format same as in Fig. 7.1.

Another check performed with the fully realistic PEs concerns the response of the analysis to MC samples whose jets have been shifted by a constant scale factor. If the set of signal MC events from which the analysis was constructed have their jets' momenta modified as  $p_j \rightarrow p_j/J$ ,  $J$  some constant, then we should expect two things. One is that if a profile of our 2-d joint likelihood is taken along the JES axis, the resulting measured JES should ideally be  $J$ , and more importantly, the other is that the measured top mass should not be affected. In other words, if our 2-d likelihood performs perfectly then, on average, the likelihoods of signal MC events should peak at the true top pole mass and at  $J$ .

PEs were run on MC signal events with five different JES values: 0.95, 0.97, 1.00, 1.03, 1.05, at three different masses,  $M_t = 167.5 \text{ GeV}/c^2$ ,  $175.0 \text{ GeV}/c^2$  and  $182.5 \text{ GeV}/c^2$ . The results are shown in Fig. 7.4, and a few conclusions can be drawn. For each of the three input masses,  $\langle M_{meas} \rangle$  is shifted at most by a few tenths of a  $\text{GeV}/c^2$  as the input JES varies; without the use of JES in the likelihood this shift would be on the order of a few  $\text{GeV}/c^2$ . Additionally, the near-unit linearity of  $\langle M_{meas} \rangle$  vs.  $M_t$  seen in Fig. 7.3 is preserved even as the input JES of the samples is changed. Finally, it is interesting to observe  $\langle \text{JES}_{meas} \rangle$  vs. input JES; here we see that the linearity response is conceivably consistent with a unity slope, but more likely a few percent less than this. As a JES measurement is not the primary aim of the analysis, this is a secondary concern.

## 7.5 Analysis Study

As a test of our final method, we ran PEs on signal and background MC in the expected fractions they'd be found in data, but with either or both of the two features described in Chap. 6 unused (the log likelihood cut and the background handling). The results of these three scenarios are shown in Table 7.1, as well as the fully realistic case in which both the log likelihood cut and the background handling are employed in the analysis. It appears that the background handling method eliminates some of the bias in the analysis, but not all; this may be due to the fact that it does not handle bad signal events. Unfortunately, it seems that the expected statistical error on the measurement at  $M_t = 172 \text{ GeV}/c^2$ , as well as the expected pull width, are virtually unchanged by the background handling technique if the likelihood cut has already been applied. The likelihood cut, on the other hand, provides significant improvements not only in the bias but also in the expected error and the pull width. Future versions of this analysis

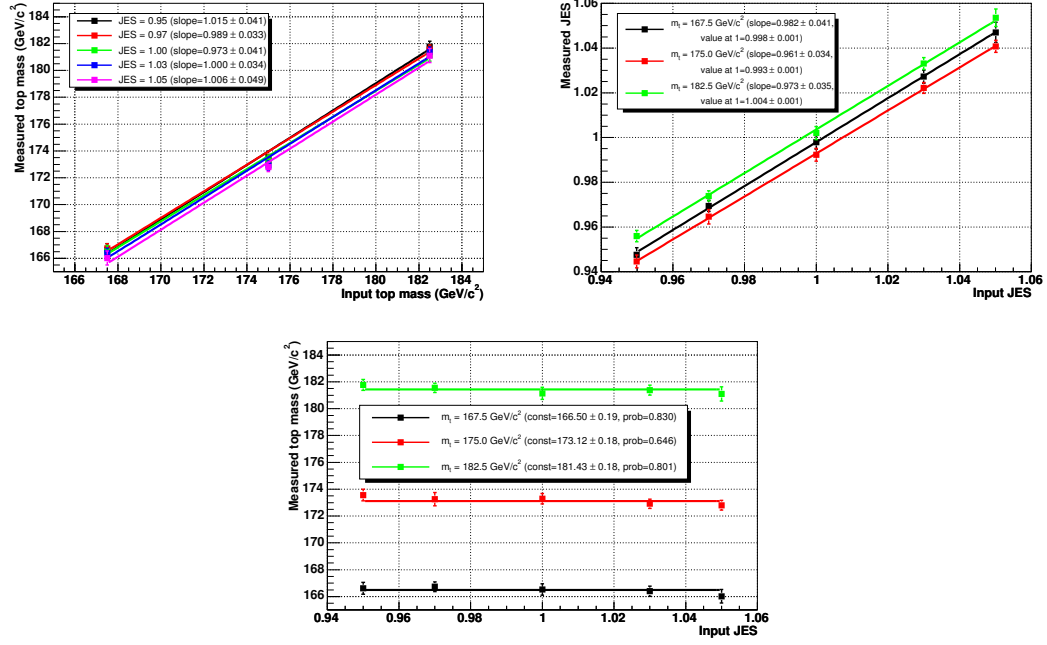


Figure 7.4: Results of the JES study. Top left,  $\langle M_{meas} \rangle$  vs.  $M_t$  for different JES values. Top right,  $\langle \text{JES}_{meas} \rangle$  vs. JES for different  $M_t$  values. Bottom,  $\langle M_{meas} \rangle$  vs. JES for different  $M_t$  values.

will attempt to improve on the manner in which undesirable events are dealt with.

Table 7.1: *PE results with the log likelihood cut and/or the background handling unused, as well as the fully realistic case in which both are used.*

Sig+Bkgnd PE Scenario	Bias	$\sigma$ at 172 GeV/c <sup>2</sup>	Pull Width
Likelihood Cut, Bkgnd Handling	$-1.20 \pm 0.14$	$2.53 \pm 0.06$	$1.22 \pm 0.02$
Likelihood Cut, No Bkgnd Handling	$-1.98 \pm 0.14$	$2.53 \pm 0.06$	$1.21 \pm 0.02$
No Likelihood Cut, Bkgnd Handling	$-2.14 \pm 0.16$	$2.83 \pm 0.06$	$1.33 \pm 0.02$
No Likelihood Cut, No Bkgnd Handling	$-4.73 \pm 0.17$	$3.01 \pm 0.06$	$1.42 \pm 0.02$



## Chapter 8

### The Measurement



## 8.1 The Measurement

The measurement presented in this thesis uses data collected by the CDF detector between February 2002 and February 2006. This data comprises 955 pb<sup>-1</sup>, out of which 149 events (108 1-tag and 41 >1-tag) passed the selection cuts described in Chap. 4 as well as the likelihood cut described in Sec. 6.2. The resulting 2-d log likelihood obtained from these events is shown in Fig. 8.1; applying the method of measurement described in Sec. 7.1, and applying the calibrations obtained from the fully realistic PEs described in Sec. 7.4, the measured top quark mass<sup>1</sup> is

$$M_{meas} = 169.75 \pm 2.28 \text{ GeV}/c^2 \quad (8.1)$$

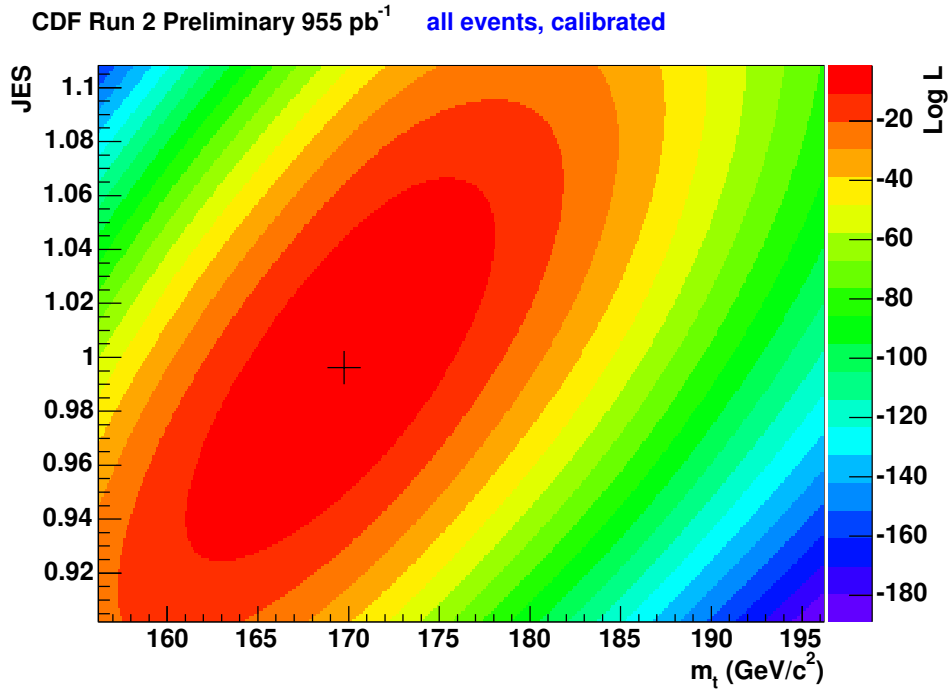


Figure 8.1: *Total 2-d log likelihood of the 149 events in the data sample.*

---

<sup>1</sup>Excluding the systematic error, described in the next Chapter

The quoted error of  $2.28 \text{ GeV}/c^2$  on the measurement is greater than the quoted error of only 27% of PEs run in the fully realistic case at  $M_t = 170.0$ ; a plot of the positive and negative errors of these PEs, as well as the positive and negative error of the data measurement, is shown in Fig. 8.2. In all cases, the errors have been corrected by the expected pull width of 1.22.

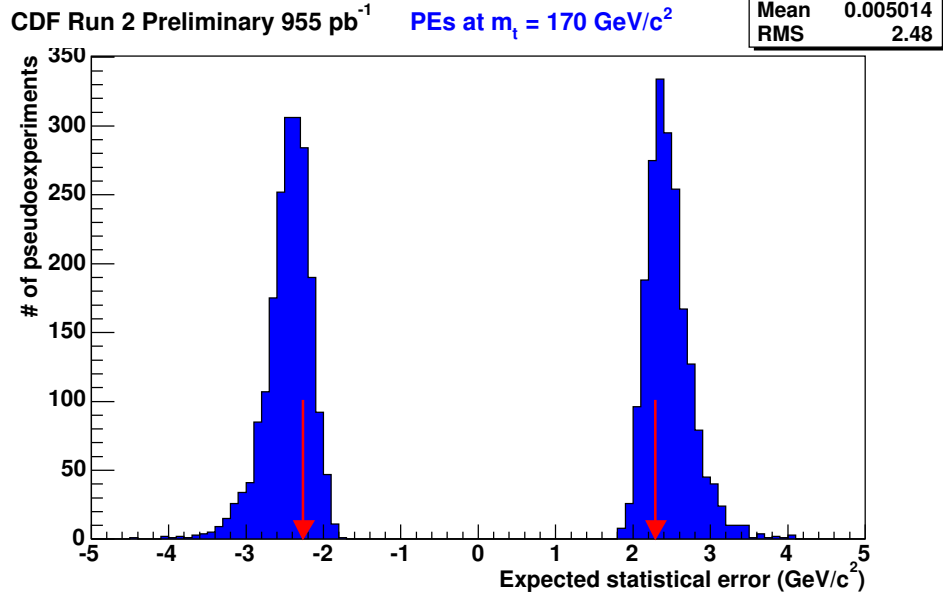


Figure 8.2: Positive and negative errors of PEs run with signal MC at  $M_t = 170.0 \text{ GeV}/c^2$ . The red arrows indicate the positive and negative errors of the data measurement.

In addition to the top mass, the JES can also be measured from the 2-d likelihood of Fig. 8.1 by taking the profile along the JES rather than the mass axis. When this is done, a JES of

$$\text{JES} = 0.996 \pm 0.017 \quad (8.2)$$

is measured. Given that the systematic uncertainty on the jet energy corrections is  $\sim 3\%$ , the measured JES is an unsurprising result. In order to obtain the JES measurement, the average of the  $\langle \text{JES}_{\text{meas}} \rangle$  vs. JES slopes at  $M_t = 167.5 \text{ GeV}/c^2$  and  $M_t = 175.0 \text{ GeV}/c^2$  from Fig. 7.4 were used to calibrate the initial

JES measurement. The quoted error of 0.017 has been scaled by a pull width of 1.14, the average of the pull widths of the JES measurement for MC samples at  $M_t = 167.5 \text{ GeV}/c^2$ , 1.17, and at  $M_t = 175.0 \text{ GeV}/c^2$ , 1.11.

## 8.2 Measurement Crosschecks

To shed further light on the measurement, top mass measurements have been performed on subsets of the total data sample; these measurements are shown in Table 8.1. The measurement performed on events with a high- $p_T$  muon only appear somewhat higher than the measurement performed on high- $E_T$  electron only events; also, the measurement performed on the  $0i$  dataset is noticeably lower than the results for the other two datasets,  $0d$  and  $0h$ . This result has been seen in other measurements on the same datasets as well [43]. Finally, it is interesting to observe that the error on the measurement using the 108 1-tag events,  $3.26 \text{ GeV}/c^2$ , is hardly any better than the error on the measurement performed using only 41  $>1$  tag events,  $3.58 \text{ GeV}/c^2$ . Reasons for this may include the higher proportion of background in the 1-tag sample compared to the  $>1$ -tag sample, as well as the possibility that the permutation weighting, described in Sec. 5.9, has more information in the case of  $>1$  tag events, only one of which had  $>2$  tags.

Finally, as a basic sanity check, after the measurement was performed a comparison was made of the both the  $\log L_{max}$  and the mass value at  $\log L_{max}$  for individual event log likelihoods between the data and MC. As can be seen in Fig. 8.3, there is excellent agreement between the data and a combination of signal and background MC events at  $M_t = 170.0 \text{ GeV}/c^2$ .

Table 8.1: *Top mass measurement summary.*

Sample	Number of Events	Measured $M_t$ (GeV/c <sup>2</sup> )
All data	149	$169.75 \pm 2.28$
electron events	88	$167.25 \pm 3.31$
muon events	61	$175.35 \pm 3.94$
1-tag events	108	$168.35 \pm 3.26$
>1-tag events	41	$171.65 \pm 3.58$
0d dataset	56	$174.25 \pm 3.55$
0h dataset	50	$171.55 \pm 4.31$
0i dataset	43	$162.25 \pm 3.96$

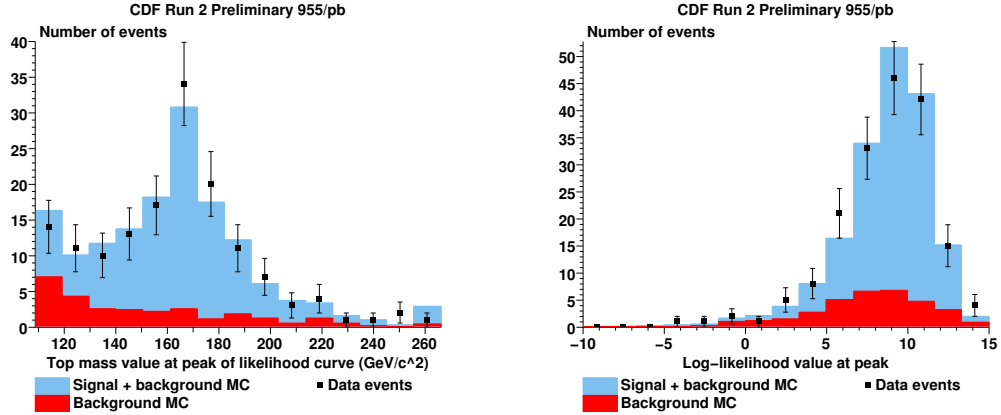


Figure 8.3: *Mass and log likelihood values at individual event likelihood maxima. Black markers are data, blue histograms are signal MC at  $M_t = 170.0$  GeV/c<sup>2</sup> and red histograms are background; signal and background are combined in their expected proportions of the data sample. Plot of masses includes likelihood cut, plot of maximum log likelihoods excludes the cut.*



# Chapter 9

## Systematics

## 9.1 Overview

Broadly speaking, there are two types of error on a measurement: statistical and systematic. The former category of error is caused by a limited amount of data, and in the limit of infinite data – e.g., an infinite number of  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  events – can be made arbitrarily small. On the other hand, some errors involve incorrect assumptions within the measurement, and therefore are *not* reduced by the amount of available data. An example of this is that the analysis is constructed and calibrated by MC events, yet it is not the case that our MC perfectly reproduces the highly complex physics of  $p\bar{p}$  collision; therefore, our MC will not perfectly mimic the types of events we have in our actual dataset and we therefore expect an error on the measurement to result.

In order to calculate a systematic error for our measurement, we typically took signal MC events which deviate from the standard MC events within the expected uncertainties of various candidate event properties. Unless otherwise noted below, Herwig events at  $M_t = 175 \text{ GeV}/c^2$  were used as well as the background events described in Sec. 4.6.3. Through performing the analysis on pseudoexperiments which use these different MC in exactly the same manner as it is performed on the data<sup>1</sup>, we can get a sense of what error our lack of perfect knowledge on various features of the physics adds to the measurement.

The individual systematic errors are assumed uncorrelated and added in quadrature to produce the total systematic of the measurement; these errors are shown in Table. 9.1 and their total is  $1.39 \text{ GeV}/c^2$ . All systematics are obtained from taking the  $\langle M_{meas} \rangle$  results of ensembles of 2000 pseudoexperiments run in the various situations. The errors quoted in the tables below are obtained using the method to calculate the error on biases described in Sec. 7.2. Most systematics were calculated using the same MC events with different ad-

---

<sup>1</sup>With the exception of some systematics we calculate specific to our analysis technique described in the next section

justments made to the aspect under study; for this reason the quoted errors for the samples used to calculate a systematic are highly correlated, and no errors on most systematics are quoted in Table 9.1. The exceptions to this are the ISR, FSR and MC Type systematics, which use completely separate MC samples in their calculation; for this reason their errors are quoted in the table.

The largest contributors to the systematic error come from our lack of knowledge as to the amount of initial and final state radiation (ISR and FSR), constituting 1.00 of the 1.39 GeV/c<sup>2</sup> error. This systematic, as well as others related to our knowledge of the background, the jet energy scale, our analysis machinery, etc., will be discussed in the following sections. In all cases, the prescriptions of the CDF Top Mass Group for systematics calculations were followed.

## 9.2 Individual Systematics

### 9.2.1 Jet Energy Scale

Inclusion of the constant scale factor “JES” into our likelihood appears to render the analysis relatively insensitive to large but constant shifts in the input jet momenta. However, in reality the systematic uncertainty to the jet energy corrections is calculated by the CDF Jet Energy Resolution group to be dependent on the  $\eta$  and  $p_T$  of the jet. To investigate the consequence of this, we ran our analysis with signal MC events whose jets’ momenta were shifted by + and -1 $\sigma$ ; the results for these samples, as well as the result obtained using the nominal Herwig MC at  $M_t = 175$  GeV/c<sup>2</sup>, are shown in Table 9.2. As the nominal reconstructed mass is higher than the reconstructed masses for either the + $\sigma$  or - $\sigma$  shifted samples, the difference between the nominal and -1  $\sigma$  result, 0.28 GeV/c<sup>2</sup>, is taken as the systematic.



Table 9.1: *Total list of systematics.*

Systematic source	Systematic error (GeV/c <sup>2</sup> )
Jet Energy Scale	0.28
$b$ Jet Energy Scale	0.23
MC Type	$0.44 \pm 0.43$
ISR	$0.75 \pm 0.36$
FSR	$0.67 \pm 0.40$
PDFs	0.46
$gg$ Fraction	0.05
Lepton $p_T$	0.05
Multiple interactions	0.05
$b$ -Tag $E_T$ Dependence	0.02
Background: Total Fraction	0.20
Background: Composition	0.39
Background: $q^2$	0.30
Calibration	0.14
Charm Tag Efficiency	0.06
$\overline{\log L_{bg}(M_t, \text{JES})}$ Shape	0.29
Total	1.39

### 9.2.2 $b$ Jet Energy Scale

The jet energy corrections used in this analysis were computed primarily with light jets; however, the systematic uncertainty on  $b$  jet energy is different than that of light jets. This is due to the different manner in which fragmentation and hadronization proceeds for the two types of jet, as well as the fact that there is uncertainty on the semileptonic decay rates of B-hadrons in  $b$  jets, a matter of concern for  $b$  but not light jets. At CDF, the  $b$ -specific systematic

Table 9.2: *Results for jet energy uncertainty.*

Sample	$\langle M_{meas} \rangle$ (GeV/c <sup>2</sup> )	$\Delta \langle M_{meas} \rangle$ (GeV/c <sup>2</sup> )
Nominal	$174.49 \pm 0.41$	—
+1 $\sigma$	$174.28 \pm 0.39$	-0.21
-1 $\sigma$	$174.21 \pm 0.4$	-0.28

has been calculated to be  $0.2 \pm 0.4$  GeV/c<sup>2</sup> due to fragmentation, 0.3% due to color flow, and 0.4% due to the semileptonic decays [48]. To account for this we took the nominal MC, and shifted the momenta of its  $b$  jets by + and -1%, resulting in  $\langle M_{meas} \rangle$  values of 174.61 and 173.83, respectively. Taking half the difference of these and multiplying by 0.6 (as per the Top Group prescription) results in a systematic of 0.23 GeV/c<sup>2</sup>.

### 9.2.3 MC Type

As mentioned earlier, this analysis is both constructed and largely tested using Herwig MC events. The question remains how accurately the MC models the  $t\bar{t}$  candidate events in data, and how much systematic error one can expect due to differences between the two types of event. One attempt to get a grasp of this is to find the shift in  $\langle M_{meas} \rangle$  when the analysis is run on Pythia signal events compared to Herwig signal events. Doing this,  $\langle M_{meas} \rangle$  is found to be  $178.70 \pm 0.22$  GeV/c<sup>2</sup> for Pythia MC at  $M_t = 178$  GeV/c<sup>2</sup>, but  $178.26 \pm 0.37$  GeV/c<sup>2</sup> for Herwig MC at  $M_t = 178$  GeV/c<sup>2</sup>. Taking the difference of the two reconstructed masses, and adding their errors in quadrature, yields a systematic of  $0.44 \pm 0.43$  GeV/c<sup>2</sup>.

### 9.2.4 ISR/FSR

Systematic uncertainty as to the amount of initial and final state radiation in  $t\bar{t}$  events has been determined at CDF from initial state radiation off of incoming quarks in Drell-Yan events [49]; this radiation was modeled by tuning parameters in Pythia-generated MC events, and as such, the systematics here are calculated using Pythia MC at  $M_t = 178 \text{ GeV}/c^2$ . The results for MC with more and less ISR and FSR are shown in Table 9.3, as well as the result for the nominal Pythia MC at  $M_t = 178 \text{ GeV}/c^2$ . In both the ISR and FSR cases, half the difference of the more and less radiation samples is less than the difference of the nominal samples with one of the radiation samples; in both cases we take the largest difference. This yields our largest systematics,  $0.75 \pm 0.36 \text{ GeV}/c^2$  for ISR and  $0.67 \pm 0.40 \text{ GeV}/c^2$  for FSR.

Table 9.3: *Systematics from ISR/FSR. All samples have a nominal  $M_t$  of 178 GeV.*

Sample	$\langle M_{meas} \rangle \text{ (GeV}/c^2)$	$\Delta \langle M_{meas} \rangle \text{ (GeV}/c^2)$
Nominal Pythia	$178.70 \pm 0.22$	—
More ISR	$178.50 \pm 0.35$	$-0.20 \pm 0.41$
Less ISR	$177.95 \pm 0.29$	$-0.75 \pm 0.36$
More FSR	$178.81 \pm 0.37$	$0.11 \pm 0.43$
Less FSR	$178.03 \pm 0.34$	$-0.67 \pm 0.40$

### 9.2.5 PDFs

A crucial aspect of modeling  $p\bar{p}$  collisions in MC events involves using Parton Distribution Functions (PDFs), functions which define the distribution of momenta among the constituent partons in a proton. Uncertainty exists on these

distributions; as a consequence we calculate a systematic using PDFs authored by different groups, PDFs with different values of  $\Lambda_{QCD}$ <sup>2</sup> and PDFs within the same model but with imperfectly known parameters of the model shifted positively and negatively. Specifically, we take the difference of PDFs authored by the MRST and CTEQ groups, MRST72 and CTEQ5L, and add them in quadrature with the difference of MRST72 ( $\Lambda_{QCD} = 228$  MeV) and MRST75 ( $\Lambda_{QCD} = 300$  MeV), as well as the difference of 20 CTEQ6M “eigenvectors”. These eigenvectors are composed of two PDFs, each PDF representing a positive or negative shift of one of the 20 parameters of the model. All of the results are plotted in Fig. 9.1; the resulting systematic is 0.46 GeV/c<sup>2</sup>.

It should be pointed out that despite having 46 sets of PDFs there were not 46 separate sets of MC generated due to the computation time this would have involved. Rather, 46 different sets of weights were applied to the nominal Pythia signal events at  $M_t = 178$  GeV/c<sup>2</sup> to effectively produce the equivalent result. It should also be noted that the method by which PEs were run in calculating this systematic were unique. The effect of statistical fluctuations on individual PDF samples’ reconstructed masses was found to dominate any systematic differences when the standard 2000 PEs were run. Therefore, for each sample, one large PE was run on all the events of the sample, with the log likelihoods of the signal events weighted according to the PDF’s set of weights and the background MC events weighted so as to recreate the usual expected fractions of the background in the sample.

As a crosscheck of these results, conventional PEs were run on independent MC for the MRST72 and MRST75 PDFs; here weighting, of course, was not used. The reconstructed masses for these two samples as well as for the nominal Pythia sample are shown in Table 9.4; the largest difference of the three, between the nominal sample and the MRST75 sample, was 0.31 GeV/c<sup>2</sup>, less than the

---

<sup>2</sup>see Sec. 2.1.2 for more on  $\Lambda_{QCD}$

calculated PDF systematic. As this was simply a sanity check, the result was not added into the total systematic for the measurement.

Table 9.4: *PDF crosscheck results*

Sample	$\langle M_{meas} \rangle \text{ GeV}/c^2$	$\Delta \langle M_{meas} \rangle \text{ GeV}/c^2$
Nominal	$178.70 \pm 0.22$	—
MRST72	$178.64 \pm 0.51$	$-0.06 \pm 0.56$
MRST75	$178.39 \pm 0.50$	$-0.31 \pm 0.55$

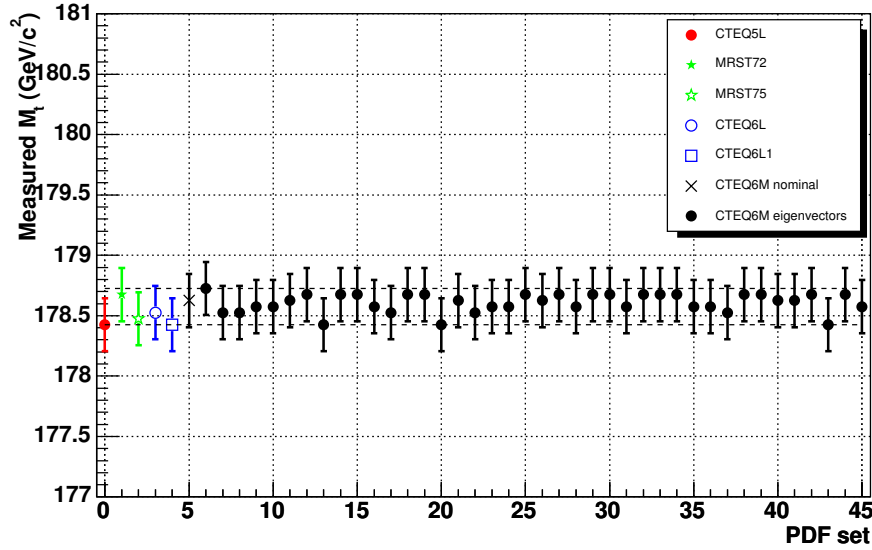


Figure 9.1: *Effect of using different PDFs on the reconstructed top mass.*

### 9.2.6 $gg$ Fraction

As the MC used in this analysis were leading-order at the generator level, only the first of the gluon fusion diagrams shown in Fig. 2.1 were used. As a consequence, gluon fusion only accounts for  $\simeq 5\%$  of  $t\bar{t}$  production in the MC, rather than the  $\simeq 15\%$  calculated at NLO [19]. A systematic is therefore

calculated by reweighting signal MC such that 15% of the events used involve gluon-gluon fusion and taking the difference with the nominal sample. As the nominal case yields  $\langle M_{meas} \rangle = 174.49 \pm 0.41 \text{ GeV}/c^2$  and the 15%  $gg$  case yields  $\langle M_{meas} \rangle = 174.44 \pm 0.40$  the difference of  $0.05 \text{ GeV}/c^2$  is taken as the systematic.

### 9.2.7 Lepton $p_T$

The systematic uncertainty on the charged lepton  $p_T$  is evaluated by taking half the difference  $\langle M_{meas} \rangle$  from PEs run on MC samples in which the lepton  $p_T$  has been shifted by  $\pm 1\%$ . When this is done, a mean reconstructed mass of  $174.50 \text{ GeV}/c^2$  is found for the increased lepton  $p_T$  and  $174.59 \text{ GeV}/c^2$  for the decreased lepton  $p_T$ , yielding a systematic of  $0.05 \text{ GeV}/c^2$ .

### 9.2.8 Multiple Interactions

In the MC used in this analysis, only one  $p\bar{p}$  interaction per bunch crossing is modeled. As a means of accounting for a systematic error due to the effects of other interactions besides  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  occurring in a bunch crossing, a systematic of  $50 \text{ MeV}/c^2$  is taken as per the Top Mass Group's directions. This decision is justified in light of the results shown in Table 9.5. PEs were run on events with different numbers of vertices from a special MC sample at  $M_t = 175 \text{ GeV}/c^2$ ; while the errors on the  $\langle M_{meas} \rangle$  values are fairly large, no large, monotonic shift in the masses is seen due to the number of interaction vertices in the event.

Table 9.5: *Study of reconstructed mass as function of number of vertices in Herwig sample at  $M_t = 175 \text{ GeV}/c^2$ .*

Sample	$\langle M_{meas} \rangle \text{ GeV}/c^2$
1 vertex	$175.72 \pm 0.69$
2 vertices	$175.83 \pm 0.76$
$> 2$ vertices	$175.53 \pm 0.52$

### 9.2.9 $b$ -Tag $E_T$ Dependence

A systematic uncertainty on the  $b$ -tagging efficiency as a function of jet  $E_T$  is modeled through two sets of weights as

$$w_+ = 0.842 + 0.00118 \cdot E_T \quad (9.1)$$

$$w_- = 0.978 - 0.00118 \cdot E_T \quad (9.2)$$

$$(9.3)$$

where here,  $E_T$  is the transverse jet energy corrected to Level 4, and  $w_+$  and  $w_-$  are  $E_T$ -dependent weights used to weight the probability for each  $b$ -tagged jet in an event. Events in the ensemble on which PEs are run are then weighted by their  $b$ -tagged jets. We find a very small effect due to this uncertainty; the  $w_+$  weights result in a mean reconstructed mass of  $174.44 \text{ GeV}/c^2$  and the  $w_-$  weights a mass of  $174.40 \text{ GeV}/c^2$ . The systematic, then is  $0.02 \text{ GeV}/c^2$ .

### 9.2.10 Background Systematics

Systematic uncertainties related to the background events in the sample concern the expected fraction of background events, the relative contributions of different types of background to the total background, and the  $q^2$  of the hard-scatter in the  $p\bar{p}$  collision which produces  $W + \text{jets}$  events.

Uncertainty on the expected fraction of background events comes from a number of sources, whose effects are added in quadrature. The errors on the expected number of background events shown in Table 4.2 are used; additionally, a 10% error on the background fraction is taken due to our extrapolation of background fractions calculated with smaller datasets to the  $955 \text{ pb}^{-1}$  dataset of this measurement. Also contributing are a 7.5% uncertainty due to potential variation of the background fraction with the jet energy scale and the binomial error resulting from our calculation of the expected number of background events (shown in Table 6.2) based on the efficiency of the likelihood cut. When the total systematic uncertainty on the background fraction is thus calculated and the background fraction of the sample is shifted up and down by this uncertainty, we obtain the results shown in Table 9.6. Half the difference of the mean reconstructed masses for the more-background and less background samples yields a systematic of  $0.20 \text{ GeV}/c^2$ .

As a (conservative) test of the effect of different proportions of background, PEs were run in which the standard background fractions were used, but the background was composed entirely of either  $W + \text{heavy}$ ,  $W + \text{light}$ , or non- $W$  events. The resulting three mean reconstructed masses are shown in Table 9.6; as the non- $W$  only PEs produced the largest difference relative to the nominal case,  $0.39 \text{ GeV}/c^2$ , this value is taken as the systematic.

A final systematic calculated based on our limited understanding of the background concerns the typical value of energy-momentum exchange,  $q^2$ , in the hard-scattering process of  $p\bar{p} \rightarrow W + \text{jets}$  events. Here we run PEs using nominal signal MC but  $W + \text{heavy}$  and  $W + \text{light}$  MC with hard-scatter  $q^2$  values equal to  $M_W^2$ ,  $(M_W/2)^2$ ,  $(2M_W)^2$ , and  $\langle p_{T_W}^2 \rangle$ , where  $\langle p_{T_W}^2 \rangle$  is the mean value of the square of the transverse momenta of the  $W$ 's in the events. The mean reconstructed masses are shown at the bottom of Table 9.6; as the largest difference between these results and the nominal sample is  $0.30 \text{ GeV}/c^2$ ,



we take this to be our systematic.

Table 9.6: *Background systematics.*

Sample	$\langle M_{meas} \rangle \text{ GeV}/c^2$	$\Delta \langle M_{meas} \rangle \text{ GeV}/c^2$
Nominal	$174.49 \pm 0.41$	—
Background Fraction +1 $\sigma$	$174.25 \pm 0.40$	-0.24
Background Fraction -1 $\sigma$	$174.65 \pm 0.38$	0.16
100% $W$ +heavy background	$174.70 \pm 0.40$	0.21
100% $W$ +light background	$174.51 \pm 0.40$	0.02
100% QCD background	$174.10 \pm 0.39$	-0.39
$q^2 = M_W^2$	$174.51 \pm 0.38$	0.02
$q^2 = 4M_W^2$	$174.20 \pm 0.38$	-0.30
$q^2 = M_W^2/4$	$174.32 \pm 0.39$	-0.17
$q^2 = \langle p_T^2 \rangle$	$174.52 \pm 0.39$	0.03

### 9.2.11 Analysis-Specific Systematics

Some of the systematic errors calculated concern uncertainties specific to our analysis. These include uncertainties on the bias calibration of the measurement, on the shape of the average background log likelihood  $\overline{\log L_{bg}(M_t, \text{JES})}$  and on the relative fraction of charm jets vs.  $b$  jets which receive  $b$  tags. The calibration systematic of  $0.14 \text{ GeV}/c^2$  is simply on the error on the bias found for the fully realistic PEs in Fig. 7.3. Additionally, while the permutation weighting used in our likelihood calculation assumes the charm jets are tagged at a rate 22% that of the  $b$  jets, there is an error of about 15% on this rate. Consequently, we run PEs in which we shift the rate from 22% up to 25.3% and down to 18.7%; these yield mean reconstructed masses of  $174.32 \text{ GeV}/c^2$  and  $174.43 \text{ GeV}/c^2$ , respectively. Half the difference of these values is our systematic,  $0.06 \text{ GeV}/c^2$ .

Another issue in the analysis concerns the systematic error introduced in the event that the  $\overline{\log L_{bg}(M_t, \text{JES})}$  we use fails to accurately model the actual  $\overline{\log L_{bg}(M_t, \text{JES})}$  in the data. In order to test this, we ran PEs where in one case we constructed our  $\overline{\log L_{bg}(M_t, \text{JES})}$  entirely from background events with a high- $p_T$  muon but used only background events with high- $E_T$  electrons in our PEs; the opposite scenario was employed as well. The results are shown in Table 9.7; we take half the difference of the resulting shifts from the two situations to yield a systematic of 0.29 GeV/c<sup>2</sup>.

Table 9.7:  $\overline{\log L_{bg}(M_t, \text{JES})}$  systematic

Sample	$< M_{meas} > \text{ GeV/c}^2$	$\Delta < M_{meas} > \text{ GeV/c}^2$
Nominal	$174.49 \pm 0.41$	—
$\mu\text{-}\overline{\log L_{bg}(M_t, \text{JES})}$ , $e$ events in PE	$174.76 \pm 0.39$	0.27
$e\text{-}\overline{\log L_{bg}(M_t, \text{JES})}$ , $\mu$ events in PE	$174.18 \pm 0.38$	-0.31

### 9.3 A Comment on 2-d vs. 1-d Likelihood

As explained in Chap. 5, a flat scale factor “JES” which multiplies the input jet momenta is added as a variable to the likelihood, turning a 1-d likelihood in  $M_t$  alone into a 2-d likelihood. The reason for this to reduce the systematic on the measurement due to the systematic uncertainties on jet energy reconstruction. As shown in Sec. 9.2.1, this systematic is calculated to be only 0.28 GeV/c<sup>2</sup>. On the other hand, what happens if only a 1-d likelihood is used,  $L(M_t, \text{JES}) \rightarrow L(M_t, \text{JES} \equiv 1)$ ? Doing this, we find that when we shift the input jets by  $+1 \sigma$  and  $-1 \sigma$  and run PEs,  $< M_{meas} >$  turns out to be 177.86 GeV/c<sup>2</sup> and 172.06 GeV/c<sup>2</sup>, respectively, resulting in a systematic of 2.90 GeV/c<sup>2</sup>! However, the 1-d likelihood, when run on the data sample, provides

a statistical error on the measurement of only  $1.6 \text{ GeV}/c^2$ , significantly lower than the quoted 2-d likelihood result of  $2.3 \text{ GeV}/c^2$ . In effect, by using a 2-d likelihood we have converted the jet energy systematic from a systematic error into a statistical error. Furthermore, the effect of the jet energy uncertainty on the total measurement is reduced using the 2-d likelihood; one can subtract off the 1-d statistical error from the 2-d error in quadrature to calculate the effect of the jet energy uncertainty in the 2-d measurement. This value,  $1.7 \text{ GeV}/c^2$ , when added in quadrature with the systematic of  $0.28 \text{ GeV}/c^2$ , is significantly lower than the 1-d systematic of  $2.90 \text{ GeV}/c^2$ . The 2-d likelihood is an improvement over the 1-d likelihood.



## Chapter 10

## Conclusions

## 10.1 Conclusions

Using 149  $t\bar{t}$  candidate events in  $955 \text{ pb}^{-1}$  of data, we have measured a top mass of  $M_{meas} = 169.8 \pm 2.3 \text{ (stat.)} \pm 1.4 \text{ (sys.) GeV}/c^2$ ; this result is consistent with the world average as of March 2007,  $M_{meas} = 170.9 \pm 1.8 \text{ GeV}/c^2$ . The technique employed was a matrix element method, a computationally intensive but increasingly popular way of measuring the top quark mass. Our approach possessed some unique features, most notably an attempt to modify the integration to account for imperfect assumptions about the kinematics as well as a background handling method which relied on a discrimination variable and knowledge of the shape of a typical background event's log likelihood.

While the analysis presented in this thesis is fundamentally sound, some new approaches will be needed to improve the method in the future. The background handling method, while capable of fixing the bias caused by background events, does not reduce the error on the measurement if the likelihood cut is also been applied. It may be that simply using a  $W + \text{jets}$  matrix element in addition to our  $t\bar{t}$  matrix elements can provide this desired improvement; all other matrix element analyses, to date, have done this. Additionally, while the log likelihood cut appears to be effective in eliminating background and bad signal events while retaining good signal events, it is a fairly crude approach; a neural network based discriminant should produce a more powerful discrimination ability. Other changes - such as an ambitious attempt to integrate over all or almost all of the 22 imperfectly known kinematic variables at the parton level of the  $t\bar{t}$  decay - may produce an improvement in the measurement. There are numerous exciting possibilities for improvement to the method which will be pursued in the future.



# Bibliography

- [1] G. Arnison *et. al*, The UA1 Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV”, Phys. Lett. B 122 (1983) 103
- [2] P. Bagnaia *et. al*, The UA2 Collaboration, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the Cern  $p\bar{p}$  Collider”, Phys. Lett. B 122 (1983) 476
- [3] G. Arnison *et. al*, The UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass around  $95 \text{ GeV}/c^2$  at the Cern SPS Collider”, Phys. Lett. B 126 (1983) 398
- [4] P. Bagnaia *et. al*, The UA2 Collaboration, “Evidence for  $Z^0 \rightarrow e^+e^-$  at the Cern  $p\bar{p}$  Collider”, Phys. Lett. B 129 (1983) 130
- [5] F. Abe *et al.*, CDF Collaboration, Phys. Rev. Lett. 74 2626 (1995);  
S. Abachi *et al.*, D0 Collaboration, Phys. Rev. Lett. 74 2632 (1995).
- [6] Lederman *et al.*, “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions,” Phys. Rev. Lett. 39 252 (1977).
- [7] K. Kondo, “Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models,” J. Phys. Soc. Jap. 57, 4126 (1988).



- [8] V. M. Abazov *et al.*, The D0 Collaboration, *Nature* (429), 638 (2004)
- [9] A. Abulencia *et al.*, The CDF Collaboration, *Phys. Rev. D* 75, 031105 (2007)
- [10] A. Abulencia *et al.*, The CDF Collaboration, FERMILAB-PUB-07-070-E
- [11] V. M. Abazov *et al.*, The D0 Collaboration, *Phys. Rev. D* 74, 092005 (2006)
- [12] W.-M. Yao *et al.*, *Journal of Physics G* 33, 1 (2006).
- [13] Francis Halzen and Alan D. Martin, Quarks and Leptons John Wiley & Sons, Inc. (1984)
- [14] Y. Fukuda *et al.*, The Super-Kamiokande Collaboration, *Phys. Rev. Lett.* 81 1562 (1998).
- [15] S. Perlmutter *et al.*, Supernova Cosmology Project Collaboration, “Measurements of Omega and Lambda from 42 High-Redshift Supernovae”, *Astrophys. J.* 517 (1999) 565.
- [16] arXiv:hep-ex/0703034v1
- [17] M. Cacciari, S. Frixione, G. Ridolfi, M.L. Mangano, and P. Nason, *JHEP* 404, 68 (2004).
- [18] D0 Note 5396-CONF
- [19] M. Cacciari *et al.*, “The t anti-t cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence”, hep-ph/0303085
- [20] I.I.Y. Bigi *et al.*, *Phys. Lett. B* 181, 157 (1986)
- [21] A. Affolder *et al.*, The CDF Collaboration, *Phys. Rev. D* 64:032002. (2001)

- [22] arXiv:hep-ph/0411179v2
- [23] A. Sill *et al.*, Nucl. Instrum. and Methods A 447, 1 (2000)
- [24] T. Affolder *et al.*, Nucl. Instrum. and Methods 526 (2004) 249.
- [25] CDF COT Group, “COT Central Outer Tracker”, CDF Note 6267, January 2003
- [26] L. Balka *et al.*, “The CDF Central Electromagnetic Calorimeter”, Nucl. Instrum. Methods, A267 (1988), p. 272.
- [27] S. Bertolucci *et al.*, “The CDF Central and Endwall Hadronic Calorimeters”, Nucl. Instrum. Methods, A267 (1988), p. 301.
- [28] R. Erbacher *et al.*, “Preliminary Event Selection and  $t\bar{t}$  Signal Acceptance for the Winter 2005 Top Lepton+Jets Sample”, CDF Note 7372, November 2004.
- [29] F. Abe *et al.*, “Topology of Three-Jet Events in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV”, Phys. Rev. D, 45 (1992), p. 1448.
- [30] A. Holloway *et al.*, “SecVtx Scale Factor Measurement for Gen5 Tagger”, CDF Note 7333, October 2004.
- [31] G. Marchesini *et al.*, Comp. Phys. Commun. 67:465, (1992) ; G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. JHEP 01:010, (2001)
- [32] M. L. Mangano, M. Moretti, G. Piccinini, R. Pittau, and A. D. Polosa, JHEP 07:001, (2003)
- [33] T. Sjostrand *et al.*, Comp. Phys. Commun. 135:238, (2001)
- [34] P. Avery, K. Read, and G. Trahem. CLEO Report CSN-212 (1985), unpublished.

- [35] S. Agostinelli *et al.*, Nucl. Instrum. Meth. A 506:250-303, (2003)
- [36] F. Canelli *et al.*, “Generic Jet Corrections for Run II”, CDF Note 7358, November 2004.
- [37] H.L.Lai *et al.*, “Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions”, EPC 12:375, (2000).
- [38] R. Kleiss and W.J.Stirling, “Top quark production at hadron colliders: some useful formulae,” ZPC 40:419 (1988).
- [39] G. Mahlon and S. Parke, “Angular Correlations in Top Quark Pair Production and Decay at Hadron Colliders”, Phys. Rev. D 53, 4886 (1996).
- [40] G. Mahlon and S. Parke, “Maximizing Spin Correlations in Top Quark Pair Production at the Tevatron”, Phys. Lett. B 411, 173 (1997).
- [41] <http://www-cdf.fnal.gov/internal/physics/top/RunIIBtag/2004/>
- [42] I. Sobol and Y.L. Levitan, “The Production of Points Uniformly Distributed in a Multidimensional Cube,” Preprint IPM Akad. Nauk SSSR, Number 40, Moscow 1976
- [43] F. Canelli and Brian Mohr, “Measurement of the Top Quark Mass with the Matrix Element Analysis Technique Using  $955 \text{ pb}^{-1}$ ”, CDF Note 8361, July 2006.
- [44] See the following books: Systems of frequency Curves by W.P. Elderton and Statistical Models in Engineering by G.J. Hahn and S.S. Shapiro.
- [45] Applied Statistics 25, 180 (1976).
- [46] S. Budd *et al.*, “Measurement of the  $t\bar{t}$  Production Cross Section in  $b$ -Tagged Lepton + Jets Events”, CDF Note 8037, January 2006.

- [47] A. Foland *et al.*, “Measurement of Top Pair Production Cross Section in Lepton + Jets Events with SecVtx Tags”, CDF Note 7536, March 2005.
- [48] J.F. Arguin *et al.*, “Combined Measurement of the Top Quark Mass and the Jet Energy Scale in the Lepton + Jets Channel”, CDF Note 7505, October 2005.
- [49] Y. K. Kim and U. K. Yang, “Initial State Gluon Radiation Studies on Drell-Yan Data for Top-Pair Production in Hadron Colliders”, CDF Note 6804, April 2004.

# Appendix A

## Data vs. MC Comparison

In order to use MC to construct and validate an analysis, it is necessary to be certain that the MC accurately models the data. For this reason, then, we created plots of various event observables, shown below, using both data and our MC, where signal at  $M_t = 175 \text{ GeV}/c^2$  and background were combined in the proportions expected given the selection cuts described in Chap. 4.

In Fig. A.1, we compare the transverse energies at correction Level 4 of the first through the fifth most energetic jets in the event. No major discrepancy is seen between the MC and the data. In Fig. A.2, we compare (clockwise from the upper left): the  $\cancel{E}_T$ , the total # of jets in the event with  $E_T > 3 \text{ GeV}$ , the total # of jets in the event with  $E_T > 8 \text{ GeV}$ , the  $p_T$  of the charged lepton and the  $E_T$  of jets which have received  $b$  tags. Here, the # of jets per event appears somewhat higher in data than in MC; fortunately this is not a primary concern in that only the four highest- $E_T$  jets are used in the analysis.

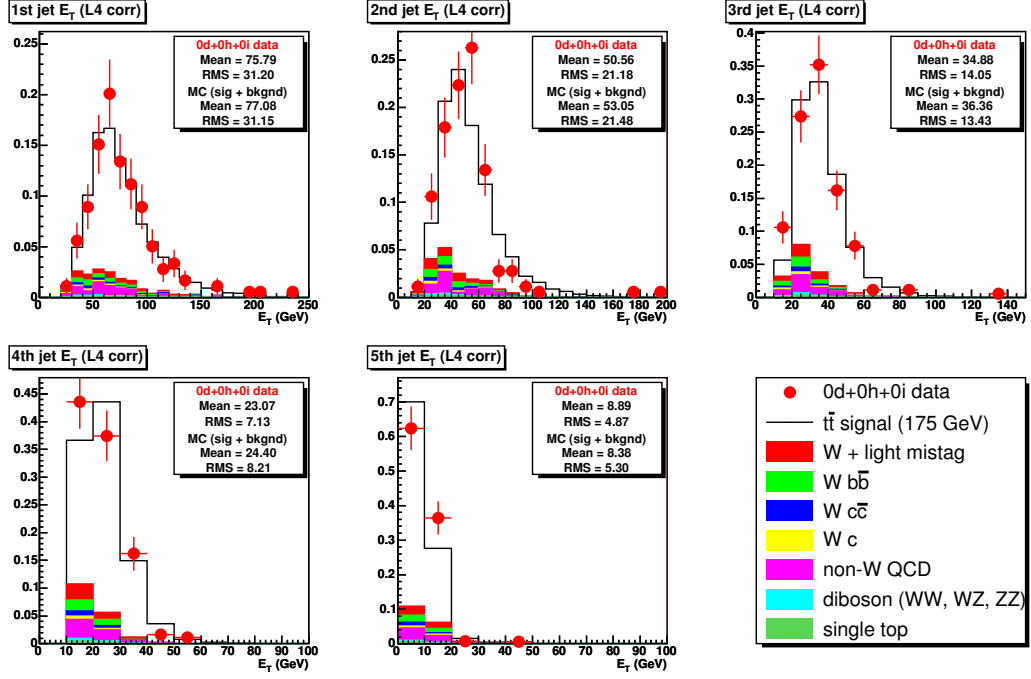


Figure A.1: Comparison of data vs. MC, jet  $E_T$ 's.

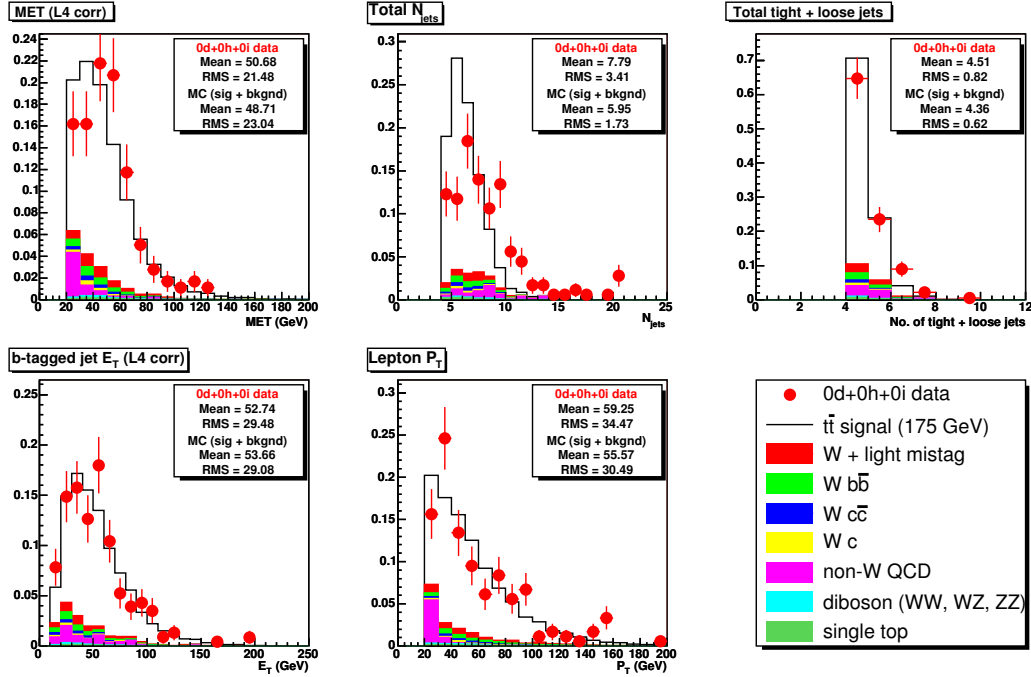


Figure A.2: Comparison of data vs. MC, various event observables.

# Appendix B

## Details of the Kleiss-Stirling Matrix Elements

In this appendix, the matrix elements developed by Kleiss and Stirling will be described and represented as spinor products; both the quark annihilation and gluon-gluon fusion processes for  $t\bar{t}$  creation, as shown in Fig. 2.1, are considered. In order to understand the algebra to follow, the following definitions of the momentum four-vectors of the incoming and outgoing products are used:

$$\bar{q}(p_1) + q(p_2) \rightarrow t(q) + \bar{t}(q') \quad (\text{B.1})$$

$$\bar{g}(p_1) + g(p_2) \rightarrow t(q) + \bar{t}(q') \quad (\text{B.2})$$

$$\bar{t}(q') \rightarrow \bar{b}(p_3) + e^-(p_4) + \bar{\nu}_e(p_5) \quad (\text{B.3})$$

$$t(q) \rightarrow \bar{b}(p_6) + \nu_e(p_7) + e^+(p_8) \quad (\text{B.4})$$

Note that the labeling of the four vectors corresponding to the different particles in the  $t\bar{t}$  production and decay is different than that found in the main body of the thesis, and that the decay products are those of the dilepton channel. The labeling reflects the same convention used in the Kleiss-Stirling paper [38], and the logic behind using dilepton decay matrix elements for  $t\bar{t} \rightarrow l\nu b\bar{b}q\bar{q}'$  events is touched on in Sec. 5.4.

It is also important to establish the notation of the products of the Dirac spinors; given  $u_+(k)$  and  $u_-(k)$  as helicity eigenstates for fermions with four-momentum  $k$ , we have:

$$\begin{aligned}s_+(k_1, k_2) &= \bar{u}_+(k_1)u_-(k_2), \\ s_-(k_1, k_2) &= \bar{u}_-(k_1)u_+(k_2),\end{aligned}\tag{B.5}$$

Finally, we shall use a function “ $P_X(k)$ ”, which is the propagator<sup>1</sup> function for particle “ $X$ ” ( $W$  and top, in practice):

$$P_X(k) = k \cdot k - M_X^2 + iM_X\Gamma_X\tag{B.6}$$

### B.0.1 Quark Annihilation Matrix Element

Let  $\nu$  denote a helicity of  $+$  or  $-$ ; temporarily postponing a discussion of the color structure, the form of the quark annihilation amplitude is:

$$M_q(\nu) = ig_s^2 g_w^4 T_q(\nu)/D_q\tag{B.7}$$

$$D_q = P_W(p_4 + p_5)P_W(p_7 + p_8)P_t(q)P_t(q')(p_1 + p_2)^2\tag{B.8}$$

here,  $g_w$  and  $g_s$  are the weak and strong coupling constants; their exponents, of course, reflect the number of vertices of their corresponding interactions in the decay.  $D_q$  involves a product of the  $W$  and top propagators, expressed in terms of their decay products, and  $T_q(\nu)$  is calculated from the spins in the process.

In order to express  $T_q(\nu)$  in terms of Dirac spinor products alone, it is necessary to introduce two new four-vectors:

$$p_9 = q' - \left( \frac{q' \cdot q'}{2q' \cdot p_4} \right) p_4\tag{B.9}$$

$$p_{10} = q - \left( \frac{q \cdot q}{2q \cdot p_8} \right) p_8\tag{B.10}$$

---

<sup>1</sup>Here, unlike in the body of the thesis, “propagator” is used in its conventional sense, rather than a term which really refers to the square of a propagator!



where  $p_9$  and  $p_{10}$  have been constructed so as to be lightlike. With these definitions, the  $T_q$  become

$$\begin{aligned}
T_q(+) &= 256 [(p_3 \cdot p_5)(p_6 \cdot p_7)]^{1/2} \\
&\quad \cdot \{-s_+(p_8, p_{10})s_-(p_{10}, p_2)s_+(p_1, p_9)s_-(p_9, p_4) + M_t^2 s_+(p_8, p_1)s_-(p_2, p_4)\}, \\
T_q(-) &= [T_q(+)](p_1 \leftrightarrow p_2)
\end{aligned} \tag{B.11}$$

With this information, the matrix elements  $M_q(+)$  and  $M_q(-)$  can be calculated, squared, and summed; a factor of 1/18 is added to account for both the color sum, and averaging over the spins and colors of the incoming quark and antiquark we have:

$$\mathcal{R}_{qq} = \frac{1}{18} \{|M_q(+)|^2 + |M_q(-)|^2\} \tag{B.12}$$

## B.0.2 Gluon-Gluon Fusion Matrix Element

As previously stated, there are three gluon fusion processes up to NLO which are accounted for by the Kleiss-Stirling matrix elements; these will be represented by  $M_g^1$ ,  $M_g^2$  and  $M_g^3$ , where each matrix element is a function of the helicities of the incoming gluons,  $\lambda_1$  and  $\lambda_2$ .  $M_g^1$  and  $M_g^2$  are the NLO processes; as  $M_g^2 = M_g^1(p_1 \leftrightarrow p_2, \lambda_1 \leftrightarrow \lambda_2)$ , only  $M_g^1$  and  $M_g^3$  will be described.

As in Eq. B.9, we need to define a new lightlike vector:

$$p_{11} = q - \left( \frac{q \cdot q}{2q \cdot p_1} \right) p_1 \tag{B.13}$$

Then the form of the gluon fusion matrix elements is:

$$\begin{aligned}
M_g^l(\lambda_1, \lambda_2) &= -ig_s^2 g_w^4 T_g^l(\lambda_1, \lambda_2) / (4D_g^l p_1 \cdot p_2), \quad l = 1, 3, \\
D_g^1 &= P_W(p_4 + p_5)P_W(p_7 + p_8)P_t(q)P_t(q')P_t(q - p_1), \\
D_g^3 &= P_W(p_4 + p_5)P_W(p_7 + p_8)P_t(q)P_t(q')(p_1 + p_2)^2.
\end{aligned} \tag{B.14}$$

and again we are tasked with calculating a matrix element which consists of

Dirac spinors in the numerator, propagators and energy terms in the denominator. For  $T_g^1$  we have:

$$T_g^1(+, +) = 512 [(p_3 \cdot p_5)(p_6 \cdot p_7)]^{1/2} \{-2(q \cdot p_1) \quad (B.15)$$

$$\begin{aligned} & \cdot s_+(p_8, p_{10}) s_-(p_{10}, p_2) s_+(p_2, p_9) s_-(p_9, p_4) \\ & + M_t^2 s_+(p_8, p_{10}) s_-(p_{10}, p_2) s_+(p_1, p_2) \\ & s_-(p_1, p_4) - M_t^2 s_+(p_8, p_1) s_-(p_2, p_1) \\ & s_+(p_2, p_9) s_-(p_9, p_4) + 2M_t^2 ((q - p_1) \cdot p_2) \\ & s_+(p_8, p_1) s_-(p_1, p_4) \}, \end{aligned}$$

$$T_g^1(+, -) = 512 [(p_3 \cdot p_5)(p_6 \cdot p_7)]^{1/2} \quad (B.16)$$

$$\begin{aligned} & \cdot \{-s_+(p_8, p_{10}) s_-(p_{10}, p_2) s_+(p_1, p_{11}) \\ & s_-(p_{11}, p_2) s_+(p_1, p_9) s_-(p_9, p_4) \\ & + M_t^2 s_+(p_8, p_1) s_-(p_2, p_{11}) s_+(p_{11}, p_1) s_-(p_2, p_4)\}, \end{aligned}$$

$$T_g^1(-, +) = [T_g^1(+, -)](p_1 \leftrightarrow p_2), \quad (B.17)$$

$$T_g^1(-, -) = 512 [(p_3 \cdot p_5)(p_6 \cdot p_7)]^{1/2} \quad (B.18)$$

$$\begin{aligned} & \cdot \{-2((q - p_1) \cdot p_2) s_+(p_8, p_{10}) s_-(p_{10}, p_1) \\ & \cdot s_+(p_2, p_9) s_-(p_9, p_4) + M_t^2 s_+(p_8, p_{10}) \\ & \cdot s_-(p_{10}, p_1) s_+(p_2, p_1) s_-(p_2, p_4) \\ & - M_t^2 s_+(p_8, p_2) s_-(p_1, p_2) s_+(p_1, p_9) s_-(p_9, p_4) \\ & 2(q \cdot p_1) s_+(p_8, p_2) s_-(p_2, p_4)\} \end{aligned}$$

and for  $T_g^3$ :

$$T_g^3(+, +) = 512 [(p_3 \cdot p_5)(p_6 \cdot p_7)]^{1/2} (p_1 \cdot p_2) \quad (\text{B.19})$$

$$\begin{aligned} & \cdot \{s_+(p_8, p_{10})s_-(p_{10}, p_1)s_+(p_1, p_9)s_-(p_9, p_4) \\ & - s_+(p_8, p_{10})s_-(p_{10}, p_2)s_+(p_2, p_9)s_-(p_9, p_4) \\ & - M_t^2(s_+(p_8, p_1)s_-(p_1, p_4) - s_+(p_8, p_2)s_-(p_2, p_4))\} , \end{aligned}$$

$$T_g^3(+, -) = T_g^3(-, +) = 0, \quad (\text{B.20})$$

$$T_g^3(-, -) = T_g^3(+, +) \quad (\text{B.21})$$

As in the case of the quark annihilation matrix element, the amplitudes are squared, and the color structure and spin configurations are also factored in, leading us to the final expression for the gluon fusion process:

$$\begin{aligned} \mathcal{R}_{gg} = \frac{1}{256} \sum_{\lambda_1, \lambda_2} & \left\{ \frac{7}{3} |M_g^1(\lambda_1, \lambda_2) + M_g^2(\lambda_1, \lambda_2)|^2 \right. \\ & \left. + 3 |M_g^1(\lambda_1, \lambda_2) - M_g^2(\lambda_1, \lambda_2) + 2M_g^3(\lambda_1, \lambda_2)|^2 \right\} . \end{aligned} \quad (\text{B.22})$$



# Appendix C

## Johnson's Functions

### C.1 Johnson's Functions

One of the advantages of using matrix element methods in top mass measurements is that, unlike in the case of a constrained fit to the kinematics of the event, it is possible to model the detector response to the energy of a particle jet with a non-gaussian function. In choosing a function for this analysis to model the distributions of jet momenta given an underlying parton  $p_T$ , then, as much flexibility as possible was desired. An ideal candidate seemed to be “Johnson’s functions” [44], a set of functions designed by the statistician Norman Johnson in 1949; these functions are capable of achieving all possible values of mean, sigma, skewness and kurtosis, where for a distribution of  $x$  with a mean of  $\mu$ , skewness is defined

$$S = \langle (x - \mu)^3 \rangle \quad (\text{C.1})$$

and likewise, for kurtosis:

$$K = \langle (x - \mu)^4 \rangle \quad (\text{C.2})$$

The Johnson's functions are defined as follows:

$$S_U : \quad p(x) = \frac{\delta}{\lambda \sqrt{2\pi} \left(1 + \left(\frac{x-\xi}{\lambda}\right)^2\right)} e^{-\frac{1}{2}(\gamma + \delta \sinh^{-1}(\frac{x-\xi}{\lambda}))^2} \quad (\text{C.3})$$

$$S_B : \quad p(x) = \frac{\delta}{\lambda \sqrt{2\pi} \left(\frac{x-\xi}{\lambda}\right) \left(1 - \frac{x-\xi}{\lambda}\right)} e^{-\frac{1}{2}(\gamma + \delta \log(\frac{x-\xi}{\xi+\lambda-x}))^2} \quad (\text{C.4})$$

where each of the two functions is defined by four parameters, separate in the case of the two functions.

$S_U$  and  $S_B$  contain some common properties of interest. The variable  $z$ , defined as  $z = \gamma + \delta f(\frac{x-\xi}{\lambda})$ , will have a normal distribution if, in the case of the  $S_u$  distribution we have it that  $f(y) = \sinh^{-1}(y)$ , or in the case of  $S_B$  we have  $f(y) = \log(\frac{y}{1-y})$ , ( $0 < y < 1$ ). Furthermore, both systems become the normal distribution as  $\delta \rightarrow \infty$  and the lognormal distribution as  $\gamma \rightarrow \infty$ .

## C.2 Fitting

While the parameters  $\delta$ ,  $\gamma$ , etc. parameterize the Johnson's functions, these are not the parameters which are directly fitted in the construction of the transfer functions. Rather, the parameters which are fit are the mean ( $\mu$ ), sigma ( $\sigma$ ), skewness ( $S$ ) and kurtosis ( $K$ ) of the distributions, where these quantities are functions of the  $p_T$  of the underlying parton of a jet as described in Sec. 5.6; we use the algorithm described in [45] to translate these to the Johnson's parameters in order to define the function.  $\mu$ ,  $\sigma$ ,  $S$  and  $K$  themselves are expressed in terms of the underlying parton  $p_T$  as  $A p_T^2 + B p_T + C$  in the case of  $\mu$ ,  $S$ , and  $K$ , and as  $D \exp(A p_T^2 + B p_T + C) + E$  in the case of  $\sigma$ , where each of the four quantities has its own set of values  $A$ ,  $B$ , etc.



# Appendix D

## Transfer Function Crosscheck

In this appendix, a derivation of the formula for the likelihood of the  $W$  mass using the transfer functions (Eq. 5.12 in Sec. 5.6.3) and the formula for the top mass likelihood (Eq. 5.13, Sec. 5.6.3) is presented.

### D.0.1 $W$ Mass Likelihood Derivation

We begin with the following formula:

$$L(\vec{p}_{j1}, \vec{p}_{j2} \mid M_W) \propto \int \frac{d^3\vec{p}_1}{2E_1} \frac{d^3\vec{p}_2}{2E_2} \delta^4(P_1 + P_2 - P_W) TF(p_{j1} \mid p_{T1}) \quad TF(p_{j2} \mid p_{T2}) F(\vec{p}_W) d^3\vec{p}_W \quad (\text{D.1})$$

where the conventions in this appendix are:

- $\vec{p}$  denotes a 3-vector.
- $p = |\vec{p}|$  denotes the magnitude of a 3-vector.
- $P$  denotes a 4-vector.
- $P^2 = P^\nu P_\nu$  denotes the invariant magnitude of a 4-vector.
- Subscripts “1” and “2” refer to the two light partons and their jets, while subscript “ $W$ ” refers to the parent  $W$



and, as in Sec. 5.6.3, constants in the likelihood are dropped, as they have no effect on the measured value. In Eq. D.1, the  $TF$ s are simply the transfer functions of the full analysis, and  $F(\vec{p}_W)$  is the prior of the momentum of the  $W$ . The parton's masses have the same assumed values as in the case of the full integration and their angles are taken to be those of their daughter jets. When (reasonably) the  $W$  momentum distribution is assumed to be independent of  $\phi$ , we can introduce the prior  $f(p_W, \cos \theta_W)$ , where, since

$$F(\vec{p}_W) d^3 \vec{p}_W = F(\vec{p}_W) p_W^2 dp_W d \cos \theta_W d\phi \quad (\text{D.2})$$

we have

$$F(\vec{p}_W) \propto f(p_W, \cos \theta_W) / p_W^2 \quad (\text{D.3})$$

assuming that the  $W$ -momenta distribution is independent of  $\phi$ . Ignoring the width of the  $W$ , we have:

$$\frac{d^3 \vec{p}_W}{2E_W} = d^4 P_W \delta(P_W^2 - M_W^2) \quad (\text{D.4})$$

and with these assumptions, Eq. D.1 becomes:

$$L(\vec{p}_{j1}, \vec{p}_{j2} \mid M_W) \propto \int \frac{p_1^2 dp_1}{E_1} \frac{p_2^2 dp_2}{E_2} \delta^4(P_1 + P_2 - P_W) TF(p_{j1} \mid p_{T1}) \\ TF(p_{j2} \mid p_{T2}) \frac{f(p_W, \cos \theta_W)}{p_W^2} 2E_W d^4 P_W \delta(P_W^2 - M_W^2) \quad (\text{D.5})$$

Now we can eliminate the  $\delta^4(P_1 + P_2 - P_W)$  in the previous equation by integrating over  $P_W$ , which gives us:

$$L(\vec{p}_{j1}, \vec{p}_{j2} \mid M_W) \propto \int \frac{p_1^2 dp_1}{E_1} \frac{p_2^2 dp_2}{E_2} TF(p_{j1} \mid p_{T1}) TF(p_{j2} \mid p_{T2}) \frac{f(p_W, \cos \theta_W)}{p_W^2} \\ 2E_W \delta(P_W^2 - M_W^2) \quad (\text{D.6})$$

There are now only two variables to be integrated on,  $p_1$  and  $p_2$ , which makes sense in light of the fact that four kinematic quantities are needed to fully describe each parton, but three are already specified by the angular and mass assumptions. Clearly, if we can transform the integration variables such that

one of them is  $P_W^2$ , then the  $\delta(P_W^2 - M_W^2)$  in the integrand can be trivially dealt with, so we select our two new variables to be  $P_W^2$  and  $\beta$ . The Jacobian for this transformation requires the derivatives of  $P_W^2$  and  $\beta$  with respect to  $p_1$  and  $p_2$ ; we observe that

$$P_W^2 = m_1^2 + m_2^2 + 2(\sqrt{p_1^2 + m_1^2}\sqrt{p_2^2 + m_2^2} - p_1 p_2 \cos \theta) \quad (\text{D.7})$$

where  $\theta$  is the angle between the two light quarks, and  $m_1$  and  $m_2$ , though assumed to be zero, are shown explicitly for the purposes of generality. So, the derivatives of  $P_W^2$  are:

$$\frac{\partial P_W^2}{\partial p_1} = 2\left(\frac{p_1 \sqrt{p_2^2 + m_2^2}}{\sqrt{p_1^2 + m_1^2}} - p_2 \cos \theta\right) = 2\left(\frac{p_1 E_2}{E_1} - p_2 \cos \theta\right) \quad (\text{D.8})$$

$$\frac{\partial P_W^2}{\partial p_2} = 2\left(\frac{p_2 \sqrt{p_1^2 + m_1^2}}{\sqrt{p_2^2 + m_2^2}} - p_1 \cos \theta\right) = 2\left(\frac{p_2 E_1}{E_2} - p_1 \cos \theta\right) \quad (\text{D.9})$$

$$\frac{\partial \beta}{\partial p_1} = \frac{1}{p_1} \quad (\text{D.10})$$

$$\frac{\partial \beta}{\partial p_2} = -\frac{1}{p_2} \quad (\text{D.11})$$

and the determinant of the resulting matrix, of course, is “ $|J_W|$ ”, the Jacobian. The final equation, then, is that shown in Eq. 5.12, reprinted here:

$$L(p_{j1}, p_{j2} \mid M_W) \propto \int \frac{p_1^2}{E_1} \frac{p_2^2}{E_2} \frac{E_W}{p_W^2} TF(p_{j1} \mid p_{T1}) TF(p_{j2} \mid p_{T2}) f(p_W, \cos \theta_W) \frac{d\beta}{|J_W|} \quad (\text{D.12})$$

## D.0.2 Top Mass Likelihood Derivation

The full formula for the integration performed to calculate the top mass is:

$$L(p_{j1}, p_{j2}, p_{j3} \mid M_t) \propto \int \frac{d^3 \vec{p}_1}{2E_1} \frac{d^3 \vec{p}_2}{2E_2} \frac{d^3 \vec{p}_b}{2E_b} \delta^4(P_1 + P_2 - P_W) TF_1 TF_2 TF_b F(\vec{p}_W) d^3 \vec{p}_W G(\vec{p}_t) \delta^4(P_W + P_b - P_t) d^3 \vec{p}_t \quad (\text{D.13})$$

where now the required terms for the top decay to the  $b$  and the  $W$  have been added, as well as the  $b$ -quark transfer function and a prior for the top

momentum,  $G(\vec{p}_t)$ . Here we use the shorthand notation  $TF_i \equiv TF(p_{ji} | p_{T_i})$ . The angle of the  $b$  quark is assumed to be the same as that of its corresponding jet, and its mass is taken to be the assumed value of  $4.95 \text{ GeV}/c^2$  as in the case of the full analysis 5.2. Simplification proceeds much as it did for the  $W$  mass case. The prior for top quark momentum is assumed to be independent of  $\phi$ ; additionally, the top quark width, like that of the  $W$ , is neglected, and the  $\delta^4(P_W + P_b - P_t)$  and  $\delta^4(P_1 + P_2 - P_W)$  are eliminated through integrating over  $P_t$  and  $P_W$ . All this yields the following:

$$L(p_{j1}, p_{j2}, p_{j3} | M_t) \propto \int \frac{p_1^2 dp_1}{E_1} \frac{p_2^2 dp_2}{E_2} \frac{p_b^2 dp_b}{E_b} TF_1 TF_2 TF_b 2E_W 2E_t \frac{f(p_W, \cos \theta_W)}{p_W^2} \frac{g(p_t, \cos \theta_t)}{p_t^2} \delta(P_W^2 - M_W^2) \delta(P_t^2 - M_t^2) \quad (\text{D.14})$$

To render the integrations over  $\delta(P_W^2 - M_W^2)$  and  $\delta(P_t^2 - M_t^2)$  trivial, we again want to transform variables, this time to  $(P_t^2, P_W^2, \beta)$  from  $(p_1, p_2, p_b)$ . Calculating the Jacobian for this transformation is quite simple as  $P_W^2$  and  $\beta$  are independent of  $p_b$ ; for this reason we have

$$|J_t| = |J_W| \cdot \frac{\partial P_t^2}{\partial p_b} \quad (\text{D.15})$$

where

$$\frac{\partial P_t^2}{\partial p_b} = 2 \left( \frac{p_b E_W}{E_b} - p_W \cos \theta_{Wb} \right) \quad (\text{D.16})$$

So, integrating out the  $\delta$ -functions  $\delta(P_W^2 - M_W^2)$  and  $\delta(P_t^2 - M_t^2)$ , the likelihood used to measure the top mass from the transfer functions is

$$L(p_{j1}, p_{j2}, p_{j3} | M_t) = \int \frac{p_1^2}{E_1} \frac{p_2^2}{E_2} \frac{p_b^2}{E_b} TF_1 TF_2 TF_b \frac{E_W}{p_W^2} \frac{E_t}{p_t^2} f(p_W, \cos \theta_W) g(p_t, \cos \theta_t) \frac{d\beta}{|J_t|} \quad (\text{D.17})$$

### D.0.3 Priors

The priors for the  $W$  and top momenta,  $f(p_W, \cos \theta_W)$  and  $g(p_t, \cos \theta_t)$ , are calculated using values taken from signal MC events which pass the selection

cuts. 2-d histograms of these values are shown in Fig. D.1; note that there's sufficiently little correlation between  $p$  and  $\cos \theta$  for both particles that we choose to model these distributions as the product of two 1-d functions. These functions are of the form  $C(x^2 + A)$  for the  $\cos \theta$  distribution of the  $W$ ,  $C(x^4 + A)$  for that of the top, and  $Cx^3 e^{-Ax}$  for the momenta distributions of both the top and the  $W$ . It should be pointed out that  $C$  and  $A$  are different for all four cases, and that  $C$  is dropped in the likelihood calculation as it is merely a constant normalization term independent of particle mass. The fits used to obtain  $C$  and  $A$  are shown for all four distributions in Fig. D.2, and can be seen to be of a high quality.

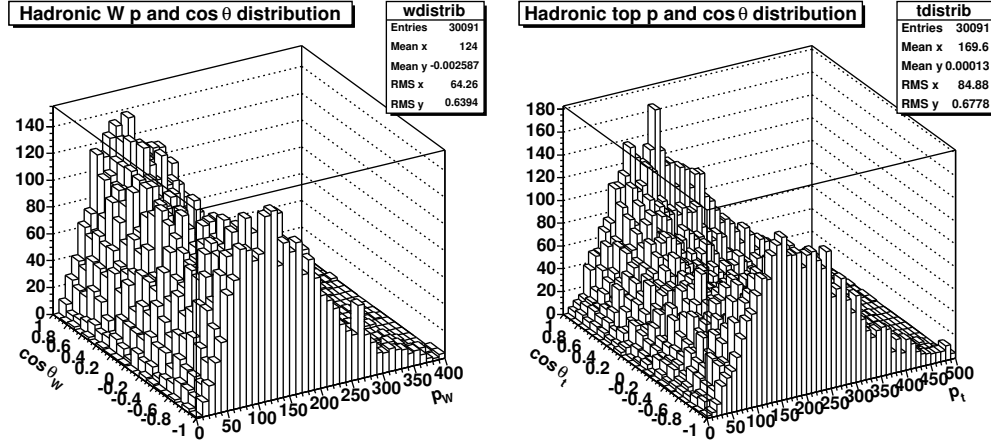


Figure D.1: 2-d distributions of  $\cos \theta$  and  $p$  taken from MC at the generator level for both the hadronic-side  $W$  and top.

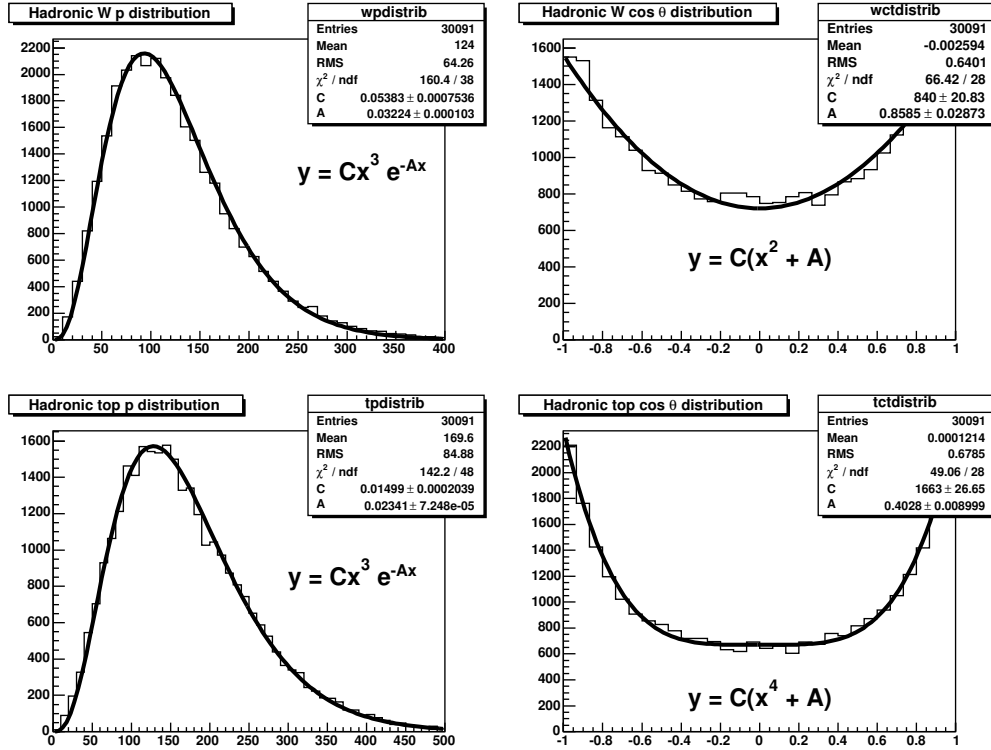


Figure D.2: 1-d functions used for the particle priors, fit to generator-level quantities.